

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 09:26:01 ; Search time 19.743 Seconds  
(without alignments)  
4694.096 Million cell updates/sec

Title: US-10-045-631B-88\_COPY\_2\_329

Perfect score: 1686

Sequence: 1 AEEGAVAVCVVRPLNSREE.....PVSFDELTALQFASTAKYM 328

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1686	100.0	2633	4	ABG06505	Abg06505 Novel hum
2	1686	100.0	2663	4	AAM39097	Aam39097 Human pol
3	1577.5	93.6	2688	4	AAM40883	Aam40883 Human pol
4	1258.5	74.6	2954	2	AYY01632	Aay01632 Amino aci
5	752	44.6	348	4	ABU53208	Abu53208 Human cel
6	752	44.6	366	4	ABU53125	Abu53125 Intracell
7	688	40.8	2013	4	ABB62332	Abb62332 Drosophil
8	643	38.1	473	5	ABG70992	Abg70992 Human tar
9	643	38.1	522	5	ABG70991	Abg70991 Human HSK
10	643	38.1	1232	7	ADD49938	Add49938 Human lun
11	643	38.1	1232	7	ADD18924	Add18924 Human dis
12	633	37.5	1232	7	ADD49937	Add49937 Human kin
13	633	37.5	1232	7	ADD49932	Add49932 Human lun
14	627.5	37.2	677	4	ABB65183	Abb65183 Drosophil
15	627	37.2	1232	5	ABG70990	Abg70990 Human HSK
16	626	37.1	1234	5	ABG70993	Abg70993 Human HSK
17	624.5	37.0	1029	5	AAE17786	Aae17786 Human kin
18	624.5	37.0	1029	7	ADC10190	Adc10190 Human NOV
19	620.5	36.8	329	5	AAE17787	Aae17787 Human kin
20	618.5	36.7	1034	3	AAE17787	Aae17787 Human kin
21	618.5	36.7	1069	3	AAE17787	Aae17787 Human kin
22	618.5	36.7	1121	3	AAE17787	Aae17787 Human kin
23	613.5	36.4	1038	5	AAE17787	Aae17787 Human kin
24	597	35.4	784	2	AAE17787	Aae17787 Human kin
25	595	35.3	1921	4	ABB62962	Abb62962 Drosophil

ALIGNMENTS

RESULT 1

ABG06505

ID ABG06505 standard; protein; 2633 AA.

XX AC

XX ABG06505;

XX AC

DT 13-FEB-2002 (first entry)

XX AC

DE Novel human diagnostic protein #6496.

XX AC

XX Human; chromosome mapping; gene mapping; forensic;

KW Human; chromosome mapping; gene mapping; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX AC

OS Homo sapiens.

XX AC

XX WO200175067-A2.

PN AC

XX AC

PD 11-OCT-2001.

XX AC

PF 30-MAR-2001; 2001WO-US008631.

XX AC

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX AC

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

 XX AC || DR | WPI; 2001-639362/73. |
DR	N-PSDB; AAS70692.
XX	AC
XX	AC
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	AC
PS	Claim 20; SEQ ID NO 36864; 103pp; English.
XX	AC
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping.
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pat\_sequences  
XX  
XX  
SQ Sequence 2633 AA;  
  
Query Match 100.0%; Score 1686; DB 4; Length 2633;  
Best Local Similarity 100.0%; Pred. No. 1.3e-155;  
Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 AEEGAVAVCVVRPLNSREESLGETAQYVWKTDDNNVYQVDGSKSFNDRVPHGNETTKN 60  
DB 2 AEEGAVAVCVVRPLNSREESLGETAQYVWKTDDNNVYQVDGSKSFNDRVPHGNETTKN 61  
  
QY 61 VVEETAAPLIIDSAIQGYNGTTFAYGQTASGKTYTMMGSEDLHGVIPRAIHDFQKIKKFP 120  
DB 62 VVEETAAPLIIDSAIQGYNGTTFAYGQTASGKTYTMMGSEDLHGVIPRAIHDFQKIKKFP 121  
  
QY 121 DREFLLRVSYMEIYNETITDLCGQTKMKPLIREDVNRNVYVADLTEEVVYTSMALKW 180  
DB 122 DREFLLRVSYMEIYNETITDLCGQTKMKPLIREDVNRNVYVADLTEEVVYTSMALKW 181  
  
QY 181 ITKGEKSRHYGETKMNQSSRSHTTFRMILSEKGEPSNCEGSKVSHNLVLDLAGSER 240  
DB 182 ITKGEKSRHYGETKMNQSSRSHTTFRMILSEKGEPSNCEGSKVSHNLVLDLAGSER 241  
  
QY 241 AAQTGAAGVRLKEGCNINRSLFILGVIKKLSDGQVGGFINYRDSKLTILQNSLGGNPK 300  
DB 242 AAQTGAAGVRLKEGCNINRSLFILGVIKKLSDGQVGGFINYRDSKLTILQNSLGGNPK 301  
  
RESULT 2  
AAM39097  
ID AAM39097 standard; protein; 2663 AA.  
AC AAM39097;  
XX  
DT 22-OCT-2001 (first entry)  
DE Human polypeptide SEQ ID NO 2242.  
XX  
XX Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
XX WO200153312-A1.  
PN  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US034263.  
XX  
XX 23-DEC-1999; 99US-00471275.  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
PR 20-JUN-2000; 2000US-00598042.  
PR 19-JUL-2000; 2000US-00620312.  
PR

PR 03-AUG-2000; 2000US-00653450.  
PR 14-SEP-2000; 2000US-00662191.  
PR 19-OCT-2000; 2000US-00693036.  
PR 29-NOV-2000; 2000US-00727344.  
XX  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
PI Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI, 2001-442253/47.  
DR N-PSDB; AAI58253.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders such  
XX as central nervous system injuries.  
PT  
PT  
XX  
XX Example 4; SEQ ID NO 2242; 10078pp; English.  
XX  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
XX encoded polypeptides (AAM38642-AAM42213) with neotropic,  
XX immunosuppressant and cytostatic activity. The polynucleotides are useful  
XX in gene therapy. A composition containing a polypeptide or polynucleotide  
XX of the invention may be used to treat diseases of the peripheral nervous  
XX system, such as peripheral nervous injuries, peripheral neuropathy and  
XX localised neuropathies and central nervous system diseases, such as  
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX utilisation of the activities such as: Immune system suppression,  
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
XX assays for receptor activity, arthritis and inflammation, leukaemias and  
XX C.N.S disorders. Note: The sequence data for this patent did not form  
XX part of the printed specification  
SQ Sequence 2663 AA;  
  
Query Match 100.0%; Score 1686; DB 4; Length 2663;  
Best Local Similarity 100.0%; Pred. No. 1.4e-155;  
Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 AEEGAVAVCVVRPLNSREESLGETAQYVWKTDDNNVYQVDGSKSFNDRVPHGNETTKN 60  
DB 2 AEEGAVAVCVVRPLNSREESLGETAQYVWKTDDNNVYQVDGSKSFNDRVPHGNETTKN 61  
  
QY 61 VVEETAAPLIIDSAIQGYNGTTFAYGQTASGKTYTMMGSEDLHGVIPRAIHDFQKIKKFP 120  
DB 62 VVEETAAPLIIDSAIQGYNGTTFAYGQTASGKTYTMMGSEDLHGVIPRAIHDFQKIKKFP 121  
  
QY 121 DREFLLRVSYMEIYNETITDLCGQTKMKPLIREDVNRNVYVADLTEEVVYTSMALKW 180  
DB 122 DREFLLRVSYMEIYNETITDLCGQTKMKPLIREDVNRNVYVADLTEEVVYTSMALKW 181  
  
QY 181 ITKGEKSRHYGETKMNQSSRSHTTFRMILSEKGEPSNCEGSKVSHNLVLDLAGSER 240  
DB 182 ITKGEKSRHYGETKMNQSSRSHTTFRMILSEKGEPSNCEGSKVSHNLVLDLAGSER 241  
  
QY 241 AAQTGAAGVRLKEGCNINRSLFILGVIKKLSDGQVGGFINYRDSKLTILQNSLGGNPK 300  
DB 242 AAQTGAAGVRLKEGCNINRSLFILGVIKKLSDGQVGGFINYRDSKLTILQNSLGGNPK 301  
  
QY 301 TRIICTITPVSFDETLTALQFASTAKYM 328  
DB 302 TRIICTITPVSFDETLTALQFASTAKYM 329  
  
RESULT 3  
AAM40883  
ID AAM40883 standard; protein; 2688 AA.  
XX  
XX AAM40883;  
AC  
XX  
DT 22-OCT-2001 (first entry)



```
Best Local Similarity 73.4%; Pred. No. 1.8e-113;
Matches 243; Conservative 40; Mismatches 43; Indels 5; Gaps 2;

QY 1 AEBGAVACVVRPLNSRELSGETAQVYWKTDNNVIYQVDGSKSFNDFRVFHNETHKN 60
Db 2 SEGDAVKCVVRPLIQREQ--GDOANLQWKAGNNTISQVDGTSFNFDRVFNSEHSTSQ 59
QY 61 VYEEIAAPIIDSALQGYNGTIFAYGOTASGKTYTMMGSEDLHGVIPRAIHDIPOKIKKFP 120
Db 60 IYOEIAPIIRLSALQGYNGTIFAYGOTSSGKTYTMMGTTPNSLGIPOAIOEVFKIIQEI 119
QY 121 DREPLRVSYMEIYNETITDLCCTQKMKPLIREDVNRNYYVADLTVEEVYTSEMALKW 180
Db 120 NREPLRVSYMEIYNETIVKOLLCDRRKKPLEIREDFNRNYYVADLTVEELVWVEHVIQW 179
QY 181 ITKGEKSRHYGETKMNQSRSHRTIFRMILESRKGEPS---NCEGSKYKSHLNLVDLAG 237
Db 180 IKKGEKSRHYGETKMNDSRSRSHRTIFRMIVESRDRNDPTNSCNDGAVMVSHLNLVDLAG 239
QY 238 SERAAOTGAAGVRLKEGCNINRSFILGOVTKKLSGQGVGFYINRDSKLTTRILONSLGG 297
Db 240 SERASQTGAEGVRLKEGCNINRSFILGOVTKKLSGQAGGFYINRDSKLTTRILONSLGG 299
QY 298 NPKTRIICTTPVSFDETLTALQFASTAKYM 328
Db 300 NAKTVIICTTPVSFDETLTSLQFASTAKHV 330

RESULT 5
ABUS3208
ID ABUS3208 standard; protein; 348 AA.
AC ABUS3208;
XX
DT 14-APR-2003 (first entry)
XX
DE Human cell cycle-associated DKFZphtes3_35b4 homologue #7.
XX
KW Human; gene therapy; vaccine; disease treatment; detection.
XX
OS Homo sapiens.
XX
FN WO200112659-A2.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-IB001496.
XX
PR 18-AUG-1999; 99US-0149499P.
PR 28-SEP-1999; 99US-0156503P.
XX
PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
PI Wiemann S;
XX
WPI; 2001-327840/34.
XX
Nucleic acids having the sequences of clones isolated from libraries of
different human tissues, useful in recombinant DNA methodologies.
XX
Example III; Page 829; 1095pp; English.
XX
This invention describes novel polynucleotides and polypeptides isolated
from human cDNA libraries which can be used for gene therapy or in
vaccines. The polynucleotides of the invention and antibodies encoded by
them may be used in the prevention, diagnosis and treatment of diseases
associated with inappropriate polypeptide expression. The products of the
invention may also be used to identify modulators of expression and
activity and to down regulate expression and activity. The antibodies of
the invention may also be used as diagnostic agents for detecting the
presence of polypeptides in samples. This sequence represents a homologue
of a polypeptide described in the disclosure of the invention
XX
```

```
SQ Sequence 348 AA;
Query Match 44.6%; Score 752; DB 4; Length 348;
Best Local Similarity 51.2%; Pred. No. 4.9e-65;
Matches 173; Conservative 36; Mismatches 103; Indels 26; Gaps 7;

QY 11 RVPLNSRELSGETAQVYV-----KTDNNVIYQVDGSKSFNDFRVFHNETHKNVYEE 64
Db 1 RCFPLNREINDGSCVQVQPPWTGYKTVHNGHEGDSPHKSFTHDFVFWNCTOEDVYDT 60
QY 65 IAAPIIDSALQGYNGTIFAYGOTASGKTYTMMG-----SEDLGVIPRAIHDIPOKIKKFP 120
Db 61 VAHPIVDCCFHYNGTIFAYGOTSGKTYTMMGGSEHDPHMGIIIPCCCHDIFDRIDKFQ 120
QY 121 DRE---FLLRVSYMEIYNETITDLC--GTQMKPLIREDVNRNYYVADLTVEEVYTSEM 176
Db 121 EKQHDFFHVKCSYMEIYNEEIIYDLCPNQMPLNIEHPNMGPIYQGTCTEPHVCSEYD 180
QY 177 ALKMITKGEKSRHYGETKMNQSRSHRTIFRMILESRKGEPSNCEGSKYKSHLNLVDLA 236
Db 181 ACHWIQGNKRNHVAATNMNDHSSRSHRTIFTHVEQRHK-----QCDEHVCCHKMNLVDLA 236
QY 237 GSEAAOTGAAGVRLKEGCNINRSFILGOVTKKLSGQV-----GGFINYRDSKLTTRI 290
Db 237 GSERVNRGTAGQRLKEGCNINQSLTTLGNVINADGQTKYMYGGHGHIFYRDSKLTWL 296
QY 291 LONSLGNGPKTRIICTTPV--SFDETLTALQFASTAK 326
Db 297 LQDSLGNCKTCMTACIWPADWNYEETLSTLYADRAK 334

RESULT 6
ABUS3125
ID ABUS3125 standard; protein; 366 AA.
AC ABUS3125;
XX
DT 15-APR-2003 (first entry)
XX
DE Intracellular trafficking-associated DKFZphtes3_26g22 homologue #2.
XX
KW Human; gene therapy; vaccine; disease treatment; detection.
XX
OS Homo sapiens.
XX
FN WO200112659-A2.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-IB001496.
XX
PR 18-AUG-1999; 99US-0149499P.
PR 28-SEP-1999; 99US-0156503P.
XX
PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
PI Wiemann S;
XX
WPI; 2001-327840/34.
XX
Nucleic acids having the sequences of clones isolated from libraries of
different human tissues, useful in recombinant DNA methodologies.
XX
Example III; Page 745; 1095pp; English.
XX
This invention describes novel polynucleotides and polypeptides isolated
from human cDNA libraries which can be used for gene therapy or in
vaccines. The polynucleotides of the invention and antibodies encoded by
them may be used in the prevention, diagnosis and treatment of diseases
associated with inappropriate polypeptide expression. The products of the
invention may also be used to identify modulators of expression and
activity and to down regulate expression and activity. The antibodies of
the invention may also be used as diagnostic agents for detecting the
presence of polypeptides in samples. This sequence represents a homologue
of a polypeptide described in the disclosure of the invention
XX
```



CC presence of polypeptides in samples. This sequence represents a homologue  
CC of a polypeptide described in the disclosure of the invention  
XX Sequence 366 AA;  
SQ

Query Match 44.6%; Score 752; DB 4; Length 366;  
Best Local Similarity 51.2%; Pred. No. 5.3e-65;  
Matches 173; Conservative 36; Mismatches 103; Indels 26; Gaps 7;  
QY 11 RVRPLNREESLGETAQYVW-----KTDNNVIVQVDSKSFNDRVPHGNETTKNVEE 64  
DB 1 RCRPLNREINDGSCVQVWPPTGYKTVNGHGEDSPHKSTFDHVFWMNCTQEDVYDT 60  
QY 65 IAPIIDSAIQYNGTIFAYGQTASGKTYTMG-----SEDLGVIPRAIHDFQKIKKFP 120  
DB 61 VAHPIDVDCFHGYNCTIFAYGQTSGKTYTMGPGGEPDHMGIIIPCCCHDIFDRIDKFQ 120  
QY 121 DRE---FLLRVSYMEIYNETITDLC-GTQMKPLIREDVNRNYYVADLTERVVTSEM 176  
DB 121 EKDHDFHWKCSYMEIYNEEYDLCNPOHMKPLNIHEHPNMGPIYVQGTETFEHVCYED 180  
QY 177 ALKWITKGEKSRHYGETKMNQSRSHRTIFRMILESREKGEPSNCEGSKVSHLNLVDLA 236  
DB 181 ACHWIWQGNKRVHVAATNMNDHSSRSHRTIFTIHEVQRHK-----QCDEHVCHSKMNLVDLA 236  
QY 237 GSERAAQTGAAGVRLKGCNINRSLFILGVQVKKLSDGV-----GGFINYRDSKLTRI 290  
DB 237 GSERVARTGAEGRLKGCNINRSLFILGVQVKKLSDGV-----GGFINYRDSKLTRI 290  
QY 291 LQNSLGNPKTRICITPV--SFDETLTALQFASTAK 326  
DB 297 LQDSLGNCKTMIACIWPADMYEETLSTIRYADRAK 334

RESULT 7  
ABB62322  
ID ABB62322 standard; protein; 2013 AA.  
XX AC ABB62322;  
XX DT 26-MAR-2002 (first entry)  
XX DE Drosophila melanogaster polypeptide SEQ ID NO 13758.  
XX KW Drosophila; developmental biology; cell signalling; insecticide;  
XX OS pharmaceutical.  
XX OS Drosophila melanogaster.  
XX PN WO200171042-A2.  
XX PD 27-SEP-2001.  
XX PF 23-MAR-2001; 2001WO-US009231.  
XX PR 23-MAR-2000; 2000US-0191637P.  
XX PR 11-JUL-2000; 2000US-00614150.  
XX PA (PEKE ) PE CORP NY.  
XX PI Venter JC, Adams M, Li PWD, Myers EW;  
XX PI WPI; 2001-656860/75.  
XX DR N-PSDB; ABL06425.  
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell  
XX PT interactions.  
XX PS Disclosure; SEQ ID NO 13758; 21pp + Sequence Listing; English.  
XX CC The invention relates to an isolated nucleic acid detection reagent  
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB116176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
CC ABBS72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 2013 AA;  
SQ

Query Match 40.8%; Score 688; DB 4; Length 2013;  
Best Local Similarity 49.8%; Pred. No. 1.3e-57;  
Matches 162; Conservative 48; Mismatches 99; Indels 16; Gaps 8;  
QY 5 AVAVCVVRPLNREESLGETAQYVWKT--DNNVIVQVDSKSFNDRVPHGNETTKNVEE 62  
DB 8 SIQVCIVKPCPEGLTSL-----WQVKERSHLADSHAEPVYFYVDEGASNOEVF 60  
QY 63 EETAAPIDSAIQYNGTIFAYGQTASGKTYTMGSEDHLGVIPRAIHDFQKIKKFPDR 122  
DB 61 DRWAKHIVHACMQGFGNTIFAYGQTSSGKTYTMGDEQNPVMVLAKEIFQIISSETER 120  
QY 123 EFLRVSYMEIYNETITDLCGTQMKPLIREDVNRNYYVADLTERVVTSEM-ALKWI 181  
DB 121 DFLRVGYIIEYNEKIYDILL--NKNQDLKHESGNGIVNVN--CECIIITSEVDLRL 176  
QY 182 TKGEKSRHYGETKMNQSRSHRTIFRMILESREKGEPSNCEGSKVSHLNLVDLAGSERA 241  
DB 177 CLGNKERTVGETNMNERSRSHAFKIIIESR-KSHSD--DDAVIOSVNLVDLAGSERA 234  
QY 242 AQTGAAGVRLKGCNINRSLFILGVQVKKLSDGVGGFINYRDSKLTRIQLNSLGGNPKT 301  
DB 235 DOTGARGARLKEGGHINKSLLFLSNVKSLSNADNRFTNYRDSKLTRIQLNSLGGNPKT 294  
QY 302 RIICITPVSFDETLTALQFASTAK 326  
DB 295 SIICITKIMSEESQSTLSFATRAK 319

RESULT 8  
ABG70992  
ID ABG70992 standard; protein; 473 AA.  
XX AC ABG70992;  
XX DT 10-DEC-2002 (first entry)  
XX DE Human target protein.  
XX KW Microtubule stimulated ATPase; cellular proliferation; cancer; AIDS;  
XX KW hyperplasia; restenosis; cardiac hypertrophy; immune disorder; HIV;  
XX KW inflammation; autoimmune disease; arthritis; graft rejection; psoriasis;  
XX KW inflammatory bowel disease; proliferation; medical procedure; surgery;  
XX KW human immunodeficiency virus; acquired immunodeficiency syndrome;  
XX KW angioplasty; human; Hskif; kinesin family.  
XX OS Homo sapiens.  
XX PN US6440684-B1.  
XX PD 27-AUG-2002.  
XX PF 12-JUN-2000; 2000US-00592054.  
XX PR 12-JUN-2000; 2000US-00592054.  
XX PA (CYTO-) CYTOKINETICS INC.  
XX PI Beraud C, Finer JT, Sakowicz R, Wood KW;  
XX PI WPI; 2002-711529/77.  
XX DR N-PSDB; ABBS55162.

human immunodeficiency virus; acquired immunodeficiency syndrome; angioplasty; human; HsKif4; kinesin family.

Homo sapiens.

Key Location/Qualifiers

Region 1..22 /note= "N-terminal T7 epitope"

Region 496..516 /note= "C-terminal myc epitope"

Region 517..522 /note= "6-histidine residues at C-terminus"

US6440684-B1.

27-AUG-2002.

12-JUN-2000; 2000US-00592054.

12-JUN-2000; 2000US-00592054.

(CYTO-) CYTOKINETICS INC.

Beraud C, Finer JT, Sakowicz R, Wood KW;

WPI: 2002-711529/77.

N-PSDB; ABS55161.

Screening for modulators of target protein having microtubule stimulated ATPase activity e.g. kinesin family of protein, useful for treating cancer, psoriasis, arthritis, human immunodeficiency virus (HIV) infection.

Claim 2; Fig 4; 34pp; English.

The present invention relates to a new method of screening modulators of target protein with microtubule stimulated ATPase activity. The method involves contacting the target protein with an agent at 1st and 2nd concentrations and determining the level of activity (e.g. binding or ATPase activity) of target protein, where a difference between levels of activity of target protein contacted with 1st and 2nd concentrations of an agent indicates that an agent modulates activity of target protein. The invention can be used for screening for modulators of target protein having microtubule stimulated ATPase activity. The compounds identified by method of the invention are useful for treating cellular proliferation including cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders and inflammation. The compounds identified by the method are also useful for treating autoimmune disease, arthritis, graft rejection, inflammatory bowel disease, proliferation induced by medical procedures, e.g. surgery, angioplasty etc. The compounds are also useful for treating psoriasis. The compounds are useful for inhibiting human immunodeficiency virus (HIV) and thus treating acquired immunodeficiency syndrome (AIDS). The present amino acid sequence represents the human HsKif4 (kinesin family) construct protein of the invention

Sequence 522 AA;

Query Match 38.1%; Score 643; DB 5; Length 522;

Best Local Similarity 44.4%; Pred. No. 4.7e-54;

Matches 147; Conservative 51; Mismatches 117; Indels 16; Gaps 5;

QY 6 VAVCVVRPLNSREESLG-ETAQVYWKTDNNVIYQVDSKSFNDRVPHGNETTKNVYEE 64  
DB VRVALRCRPLVPKEISEGQCMLSFVGPQVVGTD--KSFTYDFVDPSTEQEEVENT 87  
QY 65 IAAPTIDSAIOGNGTIFAYGQTASCKTYTMMG-----SEDHLGVIPRAIHDFQKIK 117  
DB AVAPLIKGVKFGYNATVLAYGOTSGKTSYMGAYTAEGNEPTVGVIQVLLFKEID 147  
QY 118 KFPDREFLLRVSYMEIYNETITDLCCGTQKMKPLIIREDVNRNVYVADLFEVYVSEMA 177  
DB 148 KKSDFEFTLVKSVLEIYNEEILLCLPSREKAQINIREDPKGIKIVGLTEKTVLVALDT 207

Screening for modulators of target protein having microtubule stimulated ATPase activity e.g. kinesin family of protein, useful for treating cancer, psoriasis, arthritis, human immunodeficiency virus (HIV) infection.

Claim 3; Fig 6; 34pp; English.

The present invention relates to a new method of screening modulators of target protein with microtubule stimulated ATPase activity. The method involves contacting the target protein with an agent at 1st and 2nd concentrations and determining the level of activity (e.g. binding or ATPase activity) of target protein, where a difference between levels of activity of target protein contacted with 1st and 2nd concentrations of an agent indicates that an agent modulates activity of target protein. The invention can be used for screening for modulators of target protein having microtubule stimulated ATPase activity. The compounds identified by method of the invention are useful for treating cellular proliferation including cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders and inflammation. The compounds identified by the method are also useful for treating autoimmune disease, arthritis, graft rejection, inflammatory bowel disease, proliferation induced by medical procedures, e.g. surgery, angioplasty etc. The compounds are also useful for treating psoriasis. The compounds are useful for inhibiting human immunodeficiency virus (HIV) and thus treating acquired immunodeficiency syndrome (AIDS). The present amino acid sequence represents the human target protein of the invention

Sequence 473 AA;

Query Match 38.1%; Score 643; DB 5; Length 473;

Best Local Similarity 44.4%; Pred. No. 4.1e-54;

Matches 147; Conservative 51; Mismatches 117; Indels 16; Gaps 5;

QY 6 VAVCVVRPLNSREESLG-ETAQVYWKTDNNVIYQVDSKSFNDRVPHGNETTKNVYEE 64  
DB VRVALRCRPLVPKEISEGQCMLSFVGPQVVGTD--KSFTYDFVDPSTEQEEVENT 65  
QY 8 VAVCVVRPLNSREESLG-ETAQVYWKTDNNVIYQVDSKSFNDRVPHGNETTKNVYEE 64  
DB VRVALRCRPLVPKEISEGQCMLSFVGPQVVGTD--KSFTYDFVDPSTEQEEVENT 65  
QY 65 IAAPTIDSAIOGNGTIFAYGQTASCKTYTMMG-----SEDHLGVIPRAIHDFQKIK 117  
DB AVAPLIKGVKFGYNATVLAYGOTSGKTSYMGAYTAEGNEPTVGVIQVLLFKEID 125  
QY 118 KFPDREFLLRVSYMEIYNETITDLCCGTQKMKPLIIREDVNRNVYVADLFEVYVSEMA 177  
DB 126 KKSDFEFTLVKSVLEIYNEEILLCLPSREKAQINIREDPKGIKIVGLTEKTVLVALDT 185  
QY 178 LWIKYTKGSRHYGETKNNQSRSHITIPMTLESREKGEPSNCGSVKVSHLNLDLAG 237  
DB 186 VSCLEQGNNSRTVASTAMNSQSRSRSHAITISLEORKKSD----KNSFRSKLHLVDLAG 241  
QY 238 SERAQTGAAGVRLKEGNCINRSLFTLGQVVKLSDGQVGGFVINYRDSKLTILQNSLGG 297  
DB 242 SERQKTKAEGRLKEGININRGLLCLGNVISALGDGDKGGFVYRDSKLTILQDLSLGG 301  
QY 298 NPKTRICTITPV--SFDFTLALQPASTAK 326  
DB 302 NSHTLMIACVSPADSNLEETLNTLRYADRAR 332

RESULT 9

ABG70991

ID ABG70991 standard; protein; 522 AA.

XX ABG70991;

XX 10-DEC-2002 (first entry)

XX Human HsKif4 construct protein.

XX Microtubule stimulated ATPase; cellular proliferation; cancer; AIDS;  
XX hyperplasia; restenosis; cardiac hypertrophy; immune disorder; HIV;  
XX inflammation; autoimmune disease; arthritis; graft rejection; psoriasis;  
XX inflammatory bowel disease; proliferation; medical procedure; surgery;



XX	PD	06-MAR-2003.	15-JAN-2004 (first entry)	XX	DT
XX	PF	23-AUG-2002; 2002WO-GB003892.	Human kinesin family member 4A.	XX	DE
XX	PR	23-AUG-2001; 2001GB-00020558.	Human; lung cancer antigen; cytostatic; lung cancer; gene therapy;	XX	KW
XX	PR	05-OCT-2001; 2001GB-00024037.	vaccine; T-cell; tumour.	XX	OS
XX	PA	(OXFO-) OXFORD BIOMEDICA UK LTD.	Homo sapiens.	XX	PN
XX	PI	Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;	US2003194764-A1.	XX	PD
XX	DR	N-PSDB; ADD18925.	16-OCT-2003.	XX	XX
XX	DR		04-APR-2002; 2002US-00116712.	XX	PF
XX	PT	New substantially purified polypeptide, useful for diagnosing or treating	05-APR-2001; 2001US-0282289P.	XX	XX
XX	PT	a hypoxia-regulated condition, such as cancer, ischemia, reperfusion	05-OCT-2001; 2001US-0327511P.	XX	PR
XX	PT	injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or	(CORI-) CORIXA CORP.	XX	PR
XX	PT	wound healing.		XX	XX
XX	PS	Claim 25; SEQ ID NO 413; 424pp; English.		XX	PA
XX	PS		Bangur CS, Switzer A;	XX	PI
XX	CC	This invention relates to novel human genes and gene product which are	WPI; 2003-844452/78.	XX	XX
XX	CC	implicated in certain disease states. Compounds which modulate the	N-PSDB; ADD49935.	XX	DR
XX	CC	proteins of the invention may have cytostatic, antiinflammatory, the		XX	DR
XX	CC	ophthalmological, antiarteriosclerotic or vulnerary activities. The	New isolated polypeptides and polynucleotides useful for diagnosing,	XX	PT
XX	CC	sequences of the invention may be useful for gene therapy. The invention	preventing and treating cancer, particularly lung cancer.	XX	PT
XX	CC	may be useful for diagnosing or treating a hypoxia-regulated condition,		XX	XX
XX	CC	such as tumourigenesis, angiogenesis, apoptosis, inflammation,	Example 3; SEQ ID NO 669; 250pp; English.	XX	PS
XX	CC	erythropoiesis, or the biological response to hypoxia conditions	The invention relates to an isolated polynucleotide (a) comprising any of	XX	XX
XX	CC	including processes such as glycolysis, gluconeogenesis, glucose	the 666 fully defined nucleotide sequences appearing as ADD49269 -	XX	CC
XX	CC	transportation, catecholamine synthesis, iron transport or nitric oxide	ADD49935, ADD49936 and ADD49938, complements of (a); sequences of at	XX	CC
XX	CC	synthesis. The disease includes cancer, ischaemic conditions, reperfusion	least 20 contiguous residues of (a); sequences that hybridise to (a)	XX	CC
XX	CC	injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,	under highly stringent conditions; sequences having at least 75 or 90%	XX	CC
XX	CC	inflammatory conditions or wound healing. The present sequence is that of	identity to (a); or degenerate variants of (a). Also included are an	XX	CC
XX	CC	a disease related protein of the invention.	isolated polypeptide (b) comprising: sequences encoded by the new	XX	CC
XX	XX		polynucleotide; any of the 4 amino acid sequences fully defined in the	XX	CC
XX	XX	Sequence 1232 AA;	specification; or sequences having at least 70 or 90% identity to the	XX	CC
	Query Match	38.1%; Score 643; DB 7; Length 1232;	sequence in (a) or (b)), an expression vector comprising the above	XX	CC
	Best Local Similarity	44.4%; Pred. No. 1.7e-53;	polynucleotide operably linked to an expression control sequence, a host	XX	CC
	Matches 147; Conservative 51; Mismatches 117; Indels 16; Gaps 5;		cell transformed or transfected with the above expression vector, an	XX	CC
Qy	6 VACVVRPLNSREESLG-ETAAQVWKTDNNVIYQVDSKSFNFRVFGHETTKNYYER 64		isolated antibody, or its antigen-binding fragment, that specifically	XX	CC
Db	10 VRVALRCRPLVPKEISEGCMCLSFVGPQPVVVGTD--KSFTYDFVDPSTEQEVEFNT 67		binds to the above polypeptide, an oligonucleotide that hybridises to the	XX	CC
Qy	65 IAAPLIDSAIQYNGTIFAYGOTAGSKTYWVG-----SDHLGVIPRAHIDIFQKIK 117		above-mentioned nucleotide sequences under highly stringent conditions, a	XX	CC
Db	68 AVAPLIKGVFKGYNATVLAIGTGSKTYSMGAVTAQENEPPTGVVLPVQLLFXEID 127		fusion protein comprising at least one polypeptide cited above, detecting	XX	CC
Qy	118 KPPDEFILRVSYMEIYNTIYDLCGTQMKPLAIREDVNRNVVADLVEEVYTSEMA 177		the presence of a cancer in a patient (comprising: obtaining a biological	XX	CC
Db	128 KKSDFEFLKVSLEYLYNEEILDLLCPSEKAQINIREDPKGIKIVGLTKTVLVALDT 187		sample from the patient; contacting the biological sample with a binding	XX	CC
Qy	178 LKWITKGEKSRHYGETKMNORSRSHITIFRMILESEKPEKPCGSKVSHNLVDLAG 237		agent that binds to the polypeptide, or with the oligonucleotide cited	XX	CC
Db	188 VSCLEGNNRSRTVASTAMNSQSRSRSHAITISLEQRKSD----XNSFSRSLHLVDLAG 243		above; detecting in the sample an amount of the polypeptide that binds to	XX	CC
Qy	238 SERAQGTGAAGVRLKEGCMINRSLFTLGOVKKLSDGQGVGFIVNRDLSKLTRILQNSLGG 297		the binding agent, or an amount of a polynucleotide that hybridises to	XX	CC
Db	244 SERQKTKAEGDKRLKEGININRGLLCGNVISAIGDDKKGGVFPFRDLSKLTRLLQDSLGG 303		the oligonucleotide; and comparing the amount of polypeptide, or	XX	CC
Qy	298 NPKTRICTITPV--SFDETLTALQFASTAK 326		polynucleotide that hybridises to the oligonucleotide, to a predetermined	XX	CC
Db	304 NSHTLMACVSPADSNLEETLNTLYRADRAR 334		cut-off value and then determining the presence of a cancer in the	XX	CC
	RESULT 12		patient), a method for stimulating and/or expanding T-cells specific for	XX	CC
	ADD49937		a tumour protein (comprising contacting T-cells with the above	XX	CC
XX	ID ADD49937 standard; protein; 1232 AA.		polypeptide, polynucleotide or antigen-presenting cells that express the	XX	CC
XX	AC ADD49937;		polynucleotide, under conditions and for a time sufficient to permit the	XX	CC
			stimulation and/or expansion of T-cells), an isolated T-cell population	XX	CC
			comprising T-cells prepared by the method, a composition comprising a	XX	CC
			first component selected from physiological carriers and	XX	CC
			immunostimulants, and a second component selected from the above	XX	CC
			polypeptide, polynucleotide, antibody, fusion protein, T-cell population	XX	CC
			and antigen-presenting cells that express the above polypeptide,	XX	CC
			stimulating an immune response in a patient (comprising administering to	XX	CC
			the patient the above composition) treating lung cancer in a patient	XX	CC
			(comprising administering to the patient the above composition and a	XX	CC
			diagnostic kit (comprising: at least one oligonucleotide cited above; or	XX	CC
			at least one antibody cited above and a detection reagent, where the	XX	CC
			detection reagent comprises a reporter group). The composition and	XX	CC
			methods are useful in diagnosing, preventing and treating cancer,	XX	CC
			particularly lung cancer. The present sequence is a lung cancer-	XX	CC

```
CC associated antigen of the invention.
XX SQ Sequence 1232 AA;
Query Match 37.5%; Score 633; DB 7; Length 1232;
Best Local Similarity 44.4%; Pred. No. 1.6e-52;
Matches 147; Conservative 51; Mismatches 117; Indels 16; Gaps 6;
QY 6 VAVCVVRPLNSRESLG-ETAQVYWKTDNNVIYQVDSKSFNDRVFGHNETTKNVYEE 64
DQ 10 VRVALRCRPLVPKEISEGCMCLSFVPGEPQVVGTD--KSFTYDFVDPSTQEVEFNT 67
QY 65 IAPIIDSAIOGYNGTIFAYGQTASGKTYTMWG-----SEHGLGVIPRAIHDFOKIK 117
DQ 68 AVAPLIKGVFKNATVLAAYGQTSGKTYSMGAYTAQENETPVGVIPRVQLLFKEID 127
QY 118 KFPDREFLLRVSYMEIYNETITDLLCGTQKMKPLIREDVNRNVYADLTVEVYVSEMA 177
DQ 128 KKSDFEFTLVKVSLEYNEEILDLCPREKAQINREDPKGKIVGLTEKTVLVALDT 187
QY 178 LKWITKGEKSRHYGETKMNQSRSSHTIFRMILSRKGEPSNCEGSKVSHNLVLDLAG 237
DQ 188 VSCLEQGNNSRTVASTAMNSQSSRSHAFIT---SLEQKSKDNSSFR-SKLHLVDLAG 243
QY 238 SERAAQTGAAGVRLKEGCNINRSLFILGQVKKLSGQVGGFINYRDSKLTIRLQNSLGG 297
DQ 244 SERQKTKAEGDRLEKGININRGLLCLGNVISALGDDKKGGFAPYRDSKLTIRLQDSLGG 303
QY 298 NPKTRIICITPV--SFDLTALQFATAK 326
DQ 304 NSHTLMACVSPADSNLEETLTLRYADRAR 334
RESULT 13
ID ADD49932
XX ADD49932 standard; protein; 1232 AA.
XX AC ADD49932;
XX DT 15-JAN-2004 (first entry)
XX DE Human lung specific tumour antigen L 1477p.
XX KW Human; lung cancer antigen; cytostatic; lung cancer; gene therapy;
XX KW vaccine; T-cell; tumour.
XX OS Homo sapiens.
XX PN US2003194764-A1.
XX PD 16-OCT-2003.
XX PF 04-APR-2002; 2002US-00116712.
XX PR 05-APR-2001; 2001US-0282289P.
XX PR 05-OCT-2001; 2001US-0327511P.
XX PA (CORI-) CORIXA CORP.
XX PI Bangor CS, Switzer A;
XX PI WPI; 2003-844452/78.
XX DR N-PSDB; ADD49929.
XX PT New isolated polypeptides and polynucleotides useful for diagnosing,
XX PT preventing and treating cancer, particularly lung cancer.
XX PS Example 3; SEQ ID NO 664; 250pp; English.
XX CC The invention relates to an isolated polynucleotide (a) comprising any of
XX CC the 666 fully defined nucleotide sequences appearing as ADD49269 -
XX CC ADD49935, ADD49936 and ADD49938, complements of (a); sequences of at
XX CC least 20 contiguous residues of (a); sequences that hybridise to (a)
```

```
CC under highly stringent conditions; sequences having at least 75 or 90%
CC identity to (a); or degenerate variants of (a). Also included are an
CC isolated polypeptide (b) comprising sequences encoded by the new
CC specification; any of the 4 amino acid sequences fully defined in the
CC sequence in (a) or (b), an expression vector comprising the above
CC polynucleotide operably linked to an expression control sequence, a host
CC cell transformed or transfected with the above expression vector, an
CC isolated antibody, or its antigen-binding fragment, that specifically
CC binds to the above polypeptide, an oligonucleotide that hybridises to the
CC above-mentioned nucleotide sequences under highly stringent conditions, a
CC fusion protein comprising at least one polypeptide cited above, detecting
CC the presence of a cancer in a patient (comprising: obtaining a biological
CC sample from the patient; contacting the biological sample with a binding
CC agent that binds to the polypeptide, or with the oligonucleotide cited
CC above; detecting in the sample an amount of the polypeptide that binds to
CC the binding agent, or an amount of a polynucleotide that hybridises to
CC the oligonucleotide; and comparing the amount of polypeptide, or
CC polynucleotide that hybridises to the oligonucleotide, to a predetermined
CC cut-off value and then determining the presence of a cancer in the
CC patient), a method for stimulating and/or expanding T-cells specific for
CC a tumour protein (comprising contacting T-cells with the above
CC polypeptide, polynucleotide or antigen-presenting cells that express the
CC polynucleotide, under conditions and for a time sufficient to permit the
CC stimulation and/or expansion of T-cells), an isolated T-cell population
CC comprising T-cells prepared by the method, a composition comprising a
CC first component selected from physiological carriers and
CC immunostimulants, and a second component selected from the above
CC polypeptide, polynucleotide, antibody, fusion protein, T-cell population
CC and antigen-presenting cells that express the above polypeptide,
CC stimulating an immune response in a patient (comprising administering to
CC the patient the above composition) treating lung cancer in a patient
CC (comprising administering to the patient the above composition and a
CC diagnostic kit (comprising: at least one oligonucleotide cited above; or
CC at least one antibody cited above and a detection reagent, where the
CC detection reagent comprises a reporter group). The composition and
CC methods are useful in diagnosing, preventing and treating cancer,
CC particularly lung cancer. The present sequence is a lung cancer-
CC associated antigen of the invention.
XX
SQ Sequence 1232 AA;
```

```
Query Match 37.5%; Score 633; DB 7; Length 1232;
Best Local Similarity 44.4%; Pred. No. 1.6e-52;
Matches 147; Conservative 51; Mismatches 117; Indels 16; Gaps 6;
QY 6 VAVCVVRPLNSRESLG-ETAQVYWKTDNNVIYQVDSKSFNDRVFGHNETTKNVYEE 64
DQ 10 VRVALRCRPLVPKEISEGCMCLSFVPGEPQVVGTD--KSFTYDFVDPSTQEVEFNT 67
QY 65 IAPIIDSAIOGYNGTIFAYGQTASGKTYTMWG-----SEHGLGVIPRAIHDFOKIK 117
DQ 68 AVAPLIKGVFKNATVLAAYGQTSGKTYSMGAYTAQENETPVGVIPRVQLLFKEID 127
QY 118 KFPDREFLLRVSYMEIYNETITDLLCGTQKMKPLIREDVNRNVYADLTVEVYVSEMA 177
DQ 128 KKSDFEFTLVKVSLEYNEEILDLCPREKAQINREDPKGKIVGLTEKTVLVALDT 187
QY 178 LKWITKGEKSRHYGETKMNQSRSSHTIFRMILSRKGEPSNCEGSKVSHNLVLDLAG 237
DQ 188 VSCLEQGNNSRTVASTAMNSQSSRSHAFIT---SLEQKSKDNSSFR-SKLHLVDLAG 243
QY 238 SERAAQTGAAGVRLKEGCNINRSLFILGQVKKLSGQVGGFINYRDSKLTIRLQNSLGG 297
DQ 244 SERQKTKAEGDRLEKGININRGLLCLGNVISALGDDKKGGFAPYRDSKLTIRLQDSLGG 303
QY 298 NPKTRIICITPV--SFDLTALQFATAK 326
DQ 304 NSHTLMACVSPADSNLEETLTLRYADRAR 334
RESULT 14
ABB65183
```









Result No.	Score	Query %			DB	ID	Description
		Match	Length				
1	1686	100.0	2662	4	US-09-595-684B-31	Sequence 31, Appl	
2	1258.5	74.6	2954	4	US-09-150-867-1	Sequence 1, Appl	
3	643	38.1	473	4	US-09-592-054-6	Sequence 6, Appl	
4	643	38.1	522	4	US-09-592-054-4	Sequence 4, Appl	
5	633	37.5	1231	4	US-09-595-684B-23	Sequence 23, Appl	
6	627	37.2	1232	4	US-09-592-054-2	Sequence 2, Appl	
7	626	37.1	1234	4	US-09-592-054-8	Sequence 8, Appl	
8	593.5	35.2	411	2	US-08-713-815A-4	Sequence 4, Appl	
9	593.5	35.2	441	2	US-08-713-815A-3	Sequence 3, Appl	
10	593.5	35.2	975	4	US-09-914-259-19	Sequence 19, Appl	
11	592	35.1	375	4	US-09-572-191-4	Sequence 4, Appl	
12	592	35.1	375	4	US-09-723-262-4	Sequence 4, Appl	
13	592	35.1	375	4	US-09-723-219-4	Sequence 4, Appl	
14	592	35.1	409	4	US-09-723-262-6	Sequence 6, Appl	
15	592	35.1	409	4	US-09-723-262-6	Sequence 6, Appl	
16	592	35.1	409	4	US-09-723-219-6	Sequence 6, Appl	
17	592	35.1	1388	4	US-09-572-191-2	Sequence 2, Appl	
18	592	35.1	1388	4	US-09-723-262-2	Sequence 2, Appl	
19	592	35.1	1388	4	US-09-723-219-2	Sequence 2, Appl	
20	591.5	35.1	1031	4	US-09-914-259-24	Sequence 24, Appl	
21	583	34.6	963	4	US-09-914-259-22	Sequence 22, Appl	
22	582	34.5	928	4	US-09-914-259-23	Sequence 23, Appl	
23	582	34.5	963	4	US-09-914-259-20	Sequence 20, Appl	
24	582	34.5	1066	3	US-09-541-782-8	Sequence 8, Appl	
25	582	34.5	1066	4	US-09-723-820-8	Sequence 8, Appl	
26	582	34.5	1066	4	US-10-270-085-8	Sequence 8, Appl	
27	580.5	34.4	1032	4	US-09-914-259-26	Sequence 26, Appl	

```
QY 301 TRIICITPVSPFDETLALQFASTAKYM 328
Db 302 TRIICITPVSPFDETLALQFASTAKYM 329

RESULT 2
US-09-150-867-1
; Sequence 1, Application US/09150867
; Patent No. 6645748
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth W.
; APPLICANT: Sakowicz, Roman
; APPLICANT: Goldstein, Lawrence S.B.
; APPLICANT: Cleveland, Don W.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Plus End-Directed Microtubule Motor Required for
; FILE REFERENCE: 18557C-000110US
; CURRENT APPLICATION NUMBER: US/09/150,867
; CURRENT FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: US 60/058,645
; EARLIER FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 1
; LENGTH: 2954
; TYPE: PRT
; ORGANISM: Xenopus sp.
; FEATURE:
; OTHER INFORMATION: Xenopus centromere-associated protein-B (XCENP-B)
; OTHER INFORMATION: member of the kinesin superfamily of microtubule
; OTHER INFORMATION: motor proteins
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(472)
; OTHER INFORMATION: kinesin like motor domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (473)..(2752)
; OTHER INFORMATION: rod domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (2753)..(2954)
; OTHER INFORMATION: tail domain
US-09-150-867-1

Query Match 74.6%; Score 1258.5; DB 4; Length 2954;
Best Local Similarity 73.4%; Pred. No. 6.4e-124;
Matches 243; Conservative 40; Mismatches 43; Indels 5; Gaps 2;

QY 1 APEGAVAVCVVRPLNSREESLG-ETAQVYWKTDNNVLYQVDSKSFNDRVPHGNETTKN 60
Db 2 SEGDAVKVCVRPLIQREQ--GDQANLQWKAGNNTISQVDTGKSFNDRVFNHSTSQ 59
QY 61 VYEEIAPIIDSAIOGYNGTIFAYGQTASGKTYTMMGSDHLGVIPRAIHDFQKIKKPP 120
Db 60 IYQEIAPVPIRSALQGYNGTIFAYGQTSGKTYTMMGTNSGLIIPQALQEFVKLIQETP 119
QY 121 DREFLLRVSYMEIYNETITDLCGTQKMKPLIREDVNRNVYVADLTEEVVYTSMAVKW 180
Db 120 NREFLLRVSYMEIYNETVVDLCCDRKKPLEIREDFNRNVYVADLTEELVMVPEHVIQW 179
QY 181 ITKGEKSRHYGETKMNQSRSHITFRMILESEKGEPS---NCEGSKVSHNLVLDLAG 237
Db 180 IKKGEKNRHYGETKMDHSSRSHITFRMIVESDRNDPTNSNCDGAVMVSHNLVLDLAG 239
QY 238 SERAOTGAAGVRLKEGCNINRSFILQGVIKKLSGQGVGFNYRDSKLTIRLQNSLGG 297
Db 240 SERASQTGAEGVRLKEGCNINRSFILQGVIKKLSGQAGGFNYRDSKLTIRLQNSLGG 299
QY 298 NPKTRIICTITPVSPFDETLALQFASTAKYM 328
Db 300 NAKTVIICTITPVSPFDETLALQFASTAKHV 330

us-10-045-631b-88_copy_2_329.ra1

RESULT 3
US-09-592-054-6
; Sequence 6, Application US/09592054
; Patent No. 6440684
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Finer, Jeffrey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6440684el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1016
; CURRENT APPLICATION NUMBER: US/09/592,054
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Human
US-09-592-054-6

Query Match 38.1%; Score 643; DB 4; Length 473;
Best Local Similarity 44.4%; Pred. No. 1.1e-59;
Matches 147; Conservative 51; Mismatches 117; Indels 16; Gaps 5;

QY 6 VAVCVVRPLNSREESLG-ETAQVYWKTDNNVLYQVDSKSFNDRVPHGNETTKNVYEE 64
Db 8 VRVALRCRPLVPKEISEGCMCLSFVPGEPQVVVGTD--KSFYDFVDFDPSTQEVEVNT 65
QY 65 IAAPIIDSAIOGYNGTIFAYGQTASGKTYTMMG-----SEDHLGVIPRAIHDFQKIK 117
Db 66 AVAPLIKGVFKYNATVLAGYQGTGSKTYSGGAYTAEQENEPTGVGVIPRVLIQLFKED 125
QY 118 KFPDRFLLRVSYMEIYNETITDLCGTQKMKPLIREDVNRNVYVADLTEEVVYTSMA 177
Db 126 KKSDFEFTLVKSYLEIYNEERILDLCPREKAQINREDPKGKIVGLTETKTVLVVLD 185
QY 178 LKWTIKGEKSRHYGETKMNQSRSHITFRMILESEKGEPSNCEGSKVSHNLVLDLAG 237
Db 186 VSCLEQGNNSRTVASTAMNSOSSRSHAFITISLEQRKSD---KNSSFRSKLHLVLDLAG 241
QY 238 SERAOTGAAGVRLKEGCNINRSFILQGVIKKLSGQGVGFNYRDSKLTIRLQNSLGG 297
Db 242 SERQKTKAEGDRLKEGININRGLLCLGNVISALGDKKGGFVYRDSKLTIRLQNSLGG 301
QY 298 NPKTRIICTITPV--SPFDETLALQFASTAK 326
Db 302 NSHTLMIACVSPADSNLEETLNTLEYADRAR 332

RESULT 4
US-09-592-054-4
; Sequence 4, Application US/09592054
; Patent No. 6440684
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Finer, Jeffrey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6440684el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1016
; CURRENT APPLICATION NUMBER: US/09/592,054
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Human
```

US-09-592-054-4

Query Match 38.1%; Score 643; DB 4; Length 522;  
 Best Local Similarity 44.4%; Pred. No. 1.3e-59;  
 Matches 147; Conservative 51; Mismatches 117; Indels 16; Gaps 5;

QY 6 VAVCVVRPLNSRESLG-ETAQVYWKTDNNVYQVDSKSFNDRVPHGNETTKNVEE 64  
 DB 30 VRVALRCRPLVPKEISEGCMCLSFVPGEPQVVGTD--KSFTYDFVDPSTEQEEVNT 87

QY 65 IAAPIDSAIQYNGTIFAYGQTASGKTYTMG-----SEDLGVIPRAIHDFQKIK 117  
 DB 88 AVAPLIKGVFGYNATVLAQGTGSKTYSMGAYTAQENEPTVGVIPRVIQLLFKEID 147

QY 118 KFPDREFLLRVSYMEIYNETITDLCGTQKMKPLIREDVNRNVYVADLTEEVYTSEMA 177  
 DB 148 KKSDFEFTLKVSYLEIYNEEILDLCPSREKAQINREDPKGKIVGLTEKTVLVALDT 207

QY 178 LKWTGKESRHYGETKMNQSRSHITFRMILESREKGEPSNCEGSKVSHLNLVDLAG 237  
 DB 208 VSCLEQGNNSRTVASTAMNSQSSRSHAFTI--SLEQKKSKDNSSPR-SKHLVLDLAG 243

QY 238 SERAAQTGAAGVRLKEGCNINRSIFILGOVTKKLSGQVGGFYNRDSKLTRELQNSLGG 297  
 DB 264 SERQKTKAEGDRLKEGININRGLLCGNVISALGDDKKGGFAPYRDSKLTRELQDSLGG 323

QY 298 NPKTRIITCTTPV--SFDETLTALQFASTAK 326  
 DB 324 NSHTLMIACVSPADSNLEETLNTLRYADRAR 354

## RESULT 5

US-09-595-684B-23  
 ; Sequence 23, Application US/09595684B  
 ; Patent No. 6544766

GENERAL INFORMATION:

APPLICANT: Beraud, Christophe  
 APPLICANT: Onashi, Cara  
 APPLICANT: Sakowicz, Roman  
 APPLICANT: Vaisberg, Eugeni  
 APPLICANT: Wood, Kenneth  
 APPLICANT: Yu, Ming  
 TITLE OF INVENTION: Human kinesins and methods of producing  
 TITLE OF INVENTION: and purifying human kinesins  
 FILE REFERENCE: cytop036  
 CURRENT APPLICATION NUMBER: US/09/595,684B  
 PRIOR FILING DATE: 2002-06-24  
 PRIOR APPLICATION NUMBER: 09/295,612  
 PRIOR FILING DATE: 2000-04-20  
 NUMBER OF SEQ ID NOS: 105  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 23

LENGTH: 1231

TYPE: PRT

ORGANISM: Human

US-09-595-684B-23

Query Match 37.5%; Score 633; DB 4; Length 1231;  
 Best Local Similarity 44.4%; Pred. No. 6e-58;  
 Matches 147; Conservative 51; Mismatches 117; Indels 16; Gaps 6;

QY 6 VAVCVVRPLNSRESLG-ETAQVYWKTDNNVYQVDSKSFNDRVPHGNETTKNVEE 64  
 DB 10 VRVALRCRPLVPKEISEGCMCLSFVPGEPQVVGTD--KSFTYDFVDPSTEQEEVNT 67

QY 65 IAAPIDSAIQYNGTIFAYGQTASGKTYTMG-----SEDLGVIPRAIHDFQKIK 117  
 DB 68 AVAPLIKGVFGYNATVLAQGTGSKTYSMGAYTAQENEPTVGVIPRVIQLLFKEID 127

QY 118 KFPDREFLLRVSYMEIYNETITDLCGTQKMKPLIREDVNRNVYVADLTEEVYTSEMA 177  
 DB 128 KKSDFEFTLKVSYLEIYNEEILDLCPSREKAQINREDPKGKIVGLTEKTVLVALDT 187

QY 178 LKWTGKESRHYGETKMNQSRSHITFRMILESREKGEPSNCEGSKVSHLNLVDLAG 237  
 DB 188 VSCLEQGNNSRTVASTAMNSQSSRSHAFTI--SLEQKKSKDNSSPR-SKHLVLDLAG 243

QY 238 SERAAQTGAAGVRLKEGCNINRSIFILGOVTKKLSGQVGGFYNRDSKLTRELQNSLGG 297  
 DB 244 SERQKTKAEGDRLKEGININRGLLCGNVISALGDDKKGGFAPYRDSKLTRELQDSLGG 303

QY 298 NPKTRIITCTTPV--SFDETLTALQFASTAK 326  
 DB 304 NSHTLMIACVSPADSNLEETLNTLRYADRAR 334

## RESULT 6

US-09-592-054-2  
 ; Sequence 2, Application US/09592054  
 ; Patent No. 6440684

GENERAL INFORMATION:

APPLICANT: Beraud, Christophe  
 APPLICANT: Finer, Jeffrey  
 APPLICANT: Sakowicz, Roman  
 APPLICANT: Wood, Kenneth  
 TITLE OF INVENTION: No. 6440684el motor proteins and methods for  
 TITLE OF INVENTION: their use  
 FILE REFERENCE: 1016  
 CURRENT APPLICATION NUMBER: US/09/592,054  
 CURRENT FILING DATE: 2000-07-20  
 NUMBER OF SEQ ID NOS: 8  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 2  
 LENGTH: 1232  
 TYPE: PRT  
 ORGANISM: Human  
 US-09-592-054-2

Query Match 37.2%; Score 627; DB 4; Length 1232;  
 Best Local Similarity 44.1%; Pred. No. 2.6e-57;  
 Matches 146; Conservative 51; Mismatches 118; Indels 16; Gaps 6;

QY 6 VAVCVVRPLNSRESLG-ETAQVYWKTDNNVYQVDSKSFNDRVPHGNETTKNVEE 64  
 DB 10 VRVALRCRPLVPKEISEGCMCLSFVPGEPQVVGTD--KSFTYDFVDPSTEQEEVNT 67

QY 65 IAAPIDSAIQYNGTIFAYGQTASGKTYTMG-----SEDLGVIPRAIHDFQKIK 117  
 DB 68 AVAPLIKGVFGYNATVLAQGTGSKTYSMGAYTAQENEPTVGVIPRVIQLLFKEID 127

QY 118 KFPDREFLLRVSYMEIYNETITDLCGTQKMKPLIREDVNRNVYVADLTEEVYTSEMA 177  
 DB 128 KKSDFEFTLKVSYLEIYNEEILDLCPSREKAQINREDPKGKIVGLTEKTVLVALDT 187

QY 178 LKWTGKESRHYGETKMNQSRSHITFRMILESREKGEPSNCEGSKVSHLNLVDLAG 237  
 DB 188 VSCLEQGNNSRTVASTAMNSQSSRSHAFTI--SLEQKKSKDNSSPR-SKHLVLDLAG 243

QY 238 SERAAQTGAAGVRLKEGCNINRSIFILGOVTKKLSGQVGGFYNRDSKLTRELQNSLGG 297  
 DB 244 SERQKTKAEGDRLKEGININRGLLCGNVISALGDDKKGGFAPYRDSKLTRELQDSLGG 303

QY 298 NPKTRIITCTTPV--SFDETLTALQFASTAK 326  
 DB 304 NSHTLMIACVSPADSNLEETLNTLRYADRAR 334

## RESULT 7

US-09-592-054-8  
 ; Sequence 8, Application US/09592054  
 ; Patent No. 6440684

GENERAL INFORMATION:

APPLICANT: Beraud, Christophe  
 APPLICANT: Finer, Jeffrey  
 APPLICANT: Sakowicz, Roman  
 APPLICANT: Wood, Kenneth

TELEFAX: (801)566-0750  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 411 amino acid residues  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-713-815A-4

Query Match 35.2%; Score 593.5; DB 2; Length 411;  
Best Local Similarity 45.2%; Pred. No. 1.7e-54;  
Matches 152; Conservative 42; Mismatches 119; Indels 23; Gaps 12;

QY 1 AEEGAVAVCVVRPLNSREESLG-ETAQVYKTDNNVIYQVDSKSFNFDRVHGNET 57  
Db 9 AEDSIKVV-RRPLNDEKAGKFWV--KFPNNVEENCISAG-KVILFDKVFKNAS 64  
QY 58 TKNVYBEIAPIIDSAIQYNGTIFAYGOTASCKTVM--MGSEDLGVIPRAIHDFQ 114  
Db 65 QEKVYNEAAKSIYDVLGNGTIFAYGOTSSKTHMEGVIGDSVKQGIIPRVNDIFN 124  
QY 115 KIKKEP-DREFLRVSYMEIYNETITDLCGTQRMKPLIREDVNRNYYVADLVEEVYV 173  
Db 125 HIYAMEVNFELHVKVSYEIMDKIRDL-DSKVN-LSVHEDKNRPVYKGAERFVSS 182  
QY 174 SEMALKWITKGEKSRHYGETQNRQSRSHITFRMILESREKGEPSNCEGSKVYS-HLNL 232  
Db 183 PEDVFEVIEEGKNRHLAVTNMNEHSRSHVFLNVKQ-----ENLENOKKLSGKLYL 236  
QY 233 VDLAGSRAAQTGAAGVRLKEGNNRSLFILGOVKKLSGQVGGFINYRDSKLTIRLQ 292  
Db 237 VDLAGSEKVKTAGBTGLDEAKNINKLSALGNVISALADGN-KTHIPYRDSKLTIRLQ 295  
QY 293 NSLGGPKTRIICTITPVSEF--ETLTALQFASAK 326  
Db 296 ESLGNARTTIVICSPASNESETKSLDFGRRAK 331

RESULT 9  
US-08-713-815A-3  
Sequence 3, Application US/08713815A  
Patent No. 5830659  
GENERAL INFORMATION:  
APPLICANT: Russell J. Stewart  
TITLE OF INVENTION: ACTIVE MICROTUBULE-BASED  
TITLE OF INVENTION: SEPARATIONS BY KINESINS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Thorpe, No. 5830659th & Western, L.L.P.  
STREET: 9035 South 700 East, Suite 200  
CITY: Sandy  
STATE: Utah  
COUNTRY: USA  
ZIP: 84070  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
COMPUTER: AST Ascentia 900N  
OPERATING SYSTEM: DOS 6.22  
SOFTWARE: Word Perfect 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/713,815A  
FILING DATE: 13-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Alan J. Howarth  
REGISTRATION NUMBER: 36,553  
REFERENCE/DOCKET NUMBER: T3214/U-2202  
TELEPHONE: (801)566-6633  
TELEFAX: (801)566-0750  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 441 amino acid residues  
TYPE: amino acid

TELEFAX: (801)566-0750  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 411 amino acid residues  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-713-815A-4

Query Match 35.2%; Score 593.5; DB 2; Length 411;  
Best Local Similarity 45.2%; Pred. No. 1.7e-54;  
Matches 152; Conservative 42; Mismatches 119; Indels 23; Gaps 12;

QY 1 AEEGAVAVCVVRPLNSREESLG-ETAQVYKTDNNVIYQVDSKSFNFDRVHGNET 57  
Db 9 AEDSIKVV-RRPLNDEKAGKFWV--KFPNNVEENCISAG-KVILFDKVFKNAS 64  
QY 58 TKNVYBEIAPIIDSAIQYNGTIFAYGOTASCKTVM--MGSEDLGVIPRAIHDFQ 114  
Db 65 QEKVYNEAAKSIYDVLGNGTIFAYGOTSSKTHMEGVIGDSVKQGIIPRVNDIFN 124  
QY 115 KIKKEP-DREFLRVSYMEIYNETITDLCGTQRMKPLIREDVNRNYYVADLVEEVYV 173  
Db 125 HIYAMEVNFELHVKVSYEIMDKIRDL-DSKVN-LSVHEDKNRPVYKGAERFVSS 182  
QY 174 SEMALKWITKGEKSRHYGETQNRQSRSHITFRMILESREKGEPSNCEGSKVYS-HLNL 232  
Db 183 PEDVFEVIEEGKNRHLAVTNMNEHSRSHVFLNVKQ-----ENLENOKKLSGKLYL 236  
QY 233 VDLAGSRAAQTGAAGVRLKEGNNRSLFILGOVKKLSGQVGGFINYRDSKLTIRLQ 292  
Db 237 VDLAGSEKVKTAGBTGLDEAKNINKLSALGNVISALADGN-KTHIPYRDSKLTIRLQ 295  
QY 293 NSLGGPKTRIICTITPVSEF--ETLTALQFASAK 326  
Db 296 ESLGNARTTIVICSPASNESETKSLDFGRRAK 331

RESULT 9  
US-08-713-815A-3  
Sequence 3, Application US/08713815A  
Patent No. 5830659  
GENERAL INFORMATION:  
APPLICANT: Russell J. Stewart  
TITLE OF INVENTION: ACTIVE MICROTUBULE-BASED  
TITLE OF INVENTION: SEPARATIONS BY KINESINS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Thorpe, No. 5830659th & Western, L.L.P.  
STREET: 9035 South 700 East, Suite 200  
CITY: Sandy  
STATE: Utah  
COUNTRY: USA  
ZIP: 84070  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
COMPUTER: AST Ascentia 900N  
OPERATING SYSTEM: DOS 6.22  
SOFTWARE: Word Perfect 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/713,815A  
FILING DATE: 13-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Alan J. Howarth  
REGISTRATION NUMBER: 36,553  
REFERENCE/DOCKET NUMBER: T3214/U-2202  
TELEPHONE: (801)566-6633  
TELEFAX: (801)566-0750  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 441 amino acid residues  
TYPE: amino acid



```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Human
US-09-723-262-4

Query Match      35.1%; Score 592; DB 4; Length 375;
Best Local Similarity 43.1%; Pred. No. 2.1e-54;
Matches 153; Conservative 47; Mismatches 109; Indels 46; Gaps 10;

QY 2 EGVAVVAVVVRPLNSREESL-GETAQVYWKTDNNVIYQVDS-----KSFNFD 49
Db 23 EGVAVVAVVVRPLNSREESL-GETAQVYWKTDNNVIYQVDS-----KSFNFD 73

QY 50 RVFHGNETTKNVEEIAPIIDSAIQYNGTIFAYGQTASGKTYTMWG-----SEDHL 102
Db 74 HVADVDVTTQESVFPATVAKSIVESCMSGYNGTIFAYGQTASGKTYTMWGPFSDNFHNL 133

QY 103 GVIPRAIHDF-----QKIKKFPDRFLLRVSYMEIYNETITDLCGTQKMKPLIREDV 157
Db 134 GVIPRSEYFLSLIDREKEKAGKSFCLKCSFIEIYNEQIYDLL--DSASAGLYLREHI 191

QY 158 NRNVYVADLTVEVYVTSSEMALKWITKGEKSRHYGETKNQSRSSHTIFRMILESREKGE 217
Db 192 KKGFFVVGAVEQVWTSAAEAYQVLSGGWRNRVASTSMNRRSSRSHAVFTTIESMEK-- 249

QY 218 PSNCEGSKVYSHNLNLDLAGSRAAQTGAAGVRLKEGCNINRSLFILGQVKKLSGQVG 277
Db 250 -SNEIVNIRTSLLNLDLAGSERQKDTAEGMRLKEAGNINRSLCLGQVITLVD--VG 306

QY 278 G-----FINYRDSKLTILQNSLGGNPKTRIICTITPVS--FDETILALOFASTAK 326
Db 307 NGKQRHVCYRDSKLTFLRLDSLGGNAKTAIIANVHPGSRFCFGETLSTLNFQARAK 361

RESULT 14
US-09-572-191-6
; Sequence 6, Application US/09572191
; Patent No. 6355466
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: No. 6355466el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/572,191
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Human
US-09-572-191-6

```

```

Query Match      35.1%; Score 592; DB 4; Length 409;
Best Local Similarity 43.1%; Pred. No. 2.4e-54;
Matches 153; Conservative 47; Mismatches 109; Indels 46; Gaps 10;

QY 2 EGVAVVAVVVRPLNSREESL-GETAQVYWKTDNNVIYQVDS-----KSFNFD 49
Db 23 EGVAVVAVVVRPLNSREESL-GETAQVYWKTDNNVIYQVDS-----KSFNFD 73

QY 50 RVFHGNETTKNVEEIAPIIDSAIQYNGTIFAYGQTASGKTYTMWG-----SEDHL 102
Db 74 HVADVDVTTQESVFPATVAKSIVESCMSGYNGTIFAYGQTASGKTYTMWGPFSDNFHNL 133

QY 103 GVIPRAIHDF-----QKIKKFPDRFLLRVSYMEIYNETITDLCGTQKMKPLIREDV 157
Db 134 GVIPRSEYFLSLIDREKEKAGKSFCLKCSFIEIYNEQIYDLL--DSASAGLYLREHI 191

QY 158 NRNVYVADLTVEVYVTSSEMALKWITKGEKSRHYGETKNQSRSSHTIFRMILESREKGE 217
Db 192 KKGFFVVGAVEQVWTSAAEAYQVLSGGWRNRVASTSMNRRSSRSHAVFTTIESMEK-- 249

QY 218 PSNCEGSKVYSHNLNLDLAGSRAAQTGAAGVRLKEGCNINRSLFILGQVKKLSGQVG 277
Db 250 -SNEIVNIRTSLLNLDLAGSERQKDTAEGMRLKEAGNINRSLCLGQVITLVD--VG 306

QY 278 G-----FINYRDSKLTILQNSLGGNPKTRIICTITPVS--FDETILALOFASTAK 326
Db 307 NGKQRHVCYRDSKLTFLRLDSLGGNAKTAIIANVHPGSRFCFGETLSTLNFQARAK 361

RESULT 15
US-09-723-262-6
; Sequence 6, Application US/09723262
; Patent No. 6379912
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman

```

```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Human
US-09-723-262-4

Query Match      35.1%; Score 592; DB 4; Length 375;
Best Local Similarity 43.1%; Pred. No. 2.1e-54;
Matches 153; Conservative 47; Mismatches 109; Indels 46; Gaps 10;

QY 2 EGVAVVAVVVRPLNSREESL-GETAQVYWKTDNNVIYQVDS-----KSFNFD 49
Db 23 EGVAVVAVVVRPLNSREESL-GETAQVYWKTDNNVIYQVDS-----KSFNFD 73

QY 50 RVFHGNETTKNVEEIAPIIDSAIQYNGTIFAYGQTASGKTYTMWG-----SEDHL 102
Db 74 HVADVDVTTQESVFPATVAKSIVESCMSGYNGTIFAYGQTASGKTYTMWGPFSDNFHNL 133

QY 103 GVIPRAIHDF-----QKIKKFPDRFLLRVSYMEIYNETITDLCGTQKMKPLIREDV 157
Db 134 GVIPRSEYFLSLIDREKEKAGKSFCLKCSFIEIYNEQIYDLL--DSASAGLYLREHI 191

QY 158 NRNVYVADLTVEVYVTSSEMALKWITKGEKSRHYGETKNQSRSSHTIFRMILESREKGE 217
Db 192 KKGFFVVGAVEQVWTSAAEAYQVLSGGWRNRVASTSMNRRSSRSHAVFTTIESMEK-- 249

QY 218 PSNCEGSKVYSHNLNLDLAGSRAAQTGAAGVRLKEGCNINRSLFILGQVKKLSGQVG 277
Db 250 -SNEIVNIRTSLLNLDLAGSERQKDTAEGMRLKEAGNINRSLCLGQVITLVD--VG 306

QY 278 G-----FINYRDSKLTILQNSLGGNPKTRIICTITPVS--FDETILALOFASTAK 326
Db 307 NGKQRHVCYRDSKLTFLRLDSLGGNAKTAIIANVHPGSRFCFGETLSTLNFQARAK 361

```

```

RESULT 13
US-09-723-219-4
; Sequence 4, Application US/09723219
; Patent No. 6391613
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: No. 6391613el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/723,219
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 09/572,191
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Human
US-09-723-219-4

Query Match      35.1%; Score 592; DB 4; Length 375;
Best Local Similarity 43.1%; Pred. No. 2.1e-54;
Matches 153; Conservative 47; Mismatches 109; Indels 46; Gaps 10;

QY 2 EGVAVVAVVVRPLNSREESL-GETAQVYWKTDNNVIYQVDS-----KSFNFD 49
Db 23 EGVAVVAVVVRPLNSREESL-GETAQVYWKTDNNVIYQVDS-----KSFNFD 73

QY 50 RVFHGNETTKNVEEIAPIIDSAIQYNGTIFAYGQTASGKTYTMWG-----SEDHL 102
Db 74 HVADVDVTTQESVFPATVAKSIVESCMSGYNGTIFAYGQTASGKTYTMWGPFSDNFHNL 133

QY 103 GVIPRAIHDF-----QKIKKFPDRFLLRVSYMEIYNETITDLCGTQKMKPLIREDV 157

```



```
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6379912el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/723,262
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 09/572,191
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Human
US-09-723-262-6

Query Match      35.1%; Score 592; DB 4; Length 409;
Best Local Similarity 43.1%; Pred.No.2.4e-54;
Matches 153; Conservative 47; Mismatches 109; Indels 46; Gaps 10;

QY      2 EGAVAVCVVRPINRREESL-GETAQVYWKTDNNVIYQVDGS-----KSFNFD 49
Db      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
23 EGDAIKVFVRIRPPAERSGADGE-----QNLCLSVLSSTSLRLHSNPEPKFTFD 73
QY      50 RVFHGNETTKVYEEIAAPIIDSAIOGYNGTIFAYGQTASGKTYTMWG-----SEDHL 102
Db      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
74 HVADVDTQESVFATVAKSIVSCMSGYNGTIFAYGQTSGKFTTWMGPSESDNFSHNL 133
QY      103 GVIPRAIHDF-----QKIKKFPDREFLLRVSYMEIYNETITDLICGTQWKPLIREDV 157
Db      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
134 GVIPRSFEYLFSLIDREKEKAGAGK3FLCKCFIIEYEQIYDLL--DSASAGLYLREHI 191
QY      158 NNRVYVADLTEEVYVTSEMALKWITKESRHYGETKMNQRSRSHITIFRMILESEKGE 217
Db      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
192 KKGVFVVGAVEQVWTSAAEAYQVLSGWRNRVASTMNRSSRSRSHAVFTIIESMEK-- 249
QY      218 PSNCEGSKVSHNLNVLDSERAAQTGAAGVRLKEGCNINRSLFILGQVIKKLSDGQVG 277
Db      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
250 -SNEIVNIRKTSLLNVLDSERQKDTAEGMRLKEAGNINRSLSLGQVITALVD--VG 306
QY      278 G----FINVRDSKLTRILONSLGGNPKTRIICTITPVS--FDETLTALQFASTAK 326
Db      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
307 NGQRHVVCYRDSKLTFLELSDSLGGNAKTAIIANVHPGSRFCGETLSTLNFAQRAK 361
```

Search completed: July 29, 2004, 09:42:44  
Job time : 6.81089 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: July 29, 2004, 09:39:47 ; Search time 16.3825 Seconds  
(without alignments)  
6280.361 Million cell updates/sec

Title: US-10-045-631B-88\_COPY\_2\_329  
Perfect score: 1686  
Sequence: 1 AEEGAVAVCVVRPLNSREE.....PVSFDETLALQPASTAKYM 328

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1291235 seqs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA.\*

1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	810.5	48.1	1382	16	US-10-437-963-176714
2	770.5	45.7	694	12	US-10-425-114-59725
3	664	33.4	1006	16	US-10-437-963-147933
4	651	38.6	956	16	US-10-437-963-182113
5	645.5	38.3	955	12	US-10-425-114-62672
6	643	38.1	1232	14	US-10-116-712-670
7	643	38.1	1232	15	US-10-408-765A-2153
8	643	38.1	1235	15	US-10-334-143-8
9	641.5	38.0	420	16	US-10-437-963-114373
10	638	37.8	1237	15	US-10-334-143-33
11	633	37.5	1232	14	US-10-116-712-664
12	633	37.5	1232	14	US-10-116-712-669
13	624.5	37.0	1029	16	US-10-311-642-2
14	620.5	36.8	329	16	US-10-311-642-4
15	615.5	36.5	965	16	US-10-437-963-169150

16	612.5	36.3	966	12	US-10-425-114-62697	Sequence 62697, A
17	610	36.2	672	16	US-10-408-765A-1664	Sequence 1664, Ap
18	598.5	35.5	517	15	US-10-369-493-12881	Sequence 12881, A
19	593.5	35.2	975	14	US-10-080-608A-19	Sequence 19, Appl
20	593.5	35.2	975	15	US-10-370-685-108	Sequence 108, App
21	592	35.1	375	12	US-10-332-089-4	Sequence 4, Appl
22	592	35.1	409	12	US-10-332-089-6	Sequence 6, Appl
23	592	35.1	1388	12	US-10-332-089-2	Sequence 2, Appl
24	592	35.1	1388	14	US-10-146-473-82	Sequence 82, Appl
25	592	35.1	1388	15	US-10-173-999-32	Sequence 32, Appl
26	592	35.1	1388	16	US-10-188-832-164	Sequence 164, App
27	591.5	35.1	1031	14	US-10-080-608A-24	Sequence 24, Appl
28	591.5	35.1	1031	15	US-10-370-685-113	Sequence 113, App
29	588.5	34.9	757	12	US-10-220-120-366	Sequence 366, App
30	588.5	34.9	757	16	US-10-363-829-316	Sequence 316, App
31	588	34.9	406	15	US-10-369-493-3789	Sequence 3789, Ap
32	588	34.9	1045	16	US-10-437-963-150217	Sequence 150217, A
33	584	34.6	352	12	US-10-296-838-4	Sequence 4, Appl
34	584	34.6	1362	12	US-10-296-838-2	Sequence 2, Appl
35	583.5	34.6	2552	16	US-10-437-963-129734	Sequence 129734, A
36	583	34.6	963	14	US-10-080-608A-22	Sequence 22, Appl
37	583	34.6	963	15	US-10-370-685-111	Sequence 111, App
38	582	34.5	928	14	US-10-080-608A-23	Sequence 23, Appl
39	582	34.5	928	15	US-10-370-685-112	Sequence 112, App
40	582	34.5	963	14	US-10-080-608A-20	Sequence 20, Appl
41	582	34.5	963	15	US-10-370-685-109	Sequence 109, App
42	580.5	34.4	1011	12	US-10-336-472-24	Sequence 24, Appl
43	580.5	34.4	1011	12	US-10-236-417-58	Sequence 58, Appl
44	580.5	34.4	1032	14	US-10-080-608A-26	Sequence 26, Appl
45	580.5	34.4	1032	15	US-10-370-685-115	Sequence 115, App

## ALIGNMENTS

## RESULT 1

US-10-437-963-176714  
; Sequence 176714, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 176714  
; LENGTH: 1382  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_74437C.1.pep  
US-10-437-963-176714

Query Match 48.1%; Score 810.5; DB 16; Length 1382;  
Best Local Similarity 55.0%; Pred. No. 4e-74;  
Matches 181; Conservative 43; Mismatches 88; Indels 17; Gaps 8;  
QY 6 VAVCVVRPLNSREESGETAQVWTKDNNVI-QVQDSKSFNDRVPHGNETTKQNVYE 64  
Db 4 IHVAVRPLTA--EDAGSSP---WRVSGNAIALSTQPSIRFEFRIFGEICRTADVGA 58  
QY 65 IAAPIIDSAIQVNGTIFAYGQTASGKTYMMGSEDLHGVIPRAIHDFQIKKFPREF 124  
Db 59 RTKHIVDSAVRGNGTVFAYGQNSGKTYNTRGSGNEFGIIPLAHDLFRTIEHLDREF 118

```
QY 125 LLKRVSYMEIYNETITDLGCTQKMKPLIIRVDNRRVYVADLTVEVVYVYSEMALKWITKG 184
D 119 LLMSTWMEIYNEINDLL--VPEHKLQIHESIERGLYVAGLREEIVTCPEQVLEFMSFG 176
QY 185 EKSRRHYGETKQNRSSRSHITFRMILSRKGPSS-----NCEGSVKVSHLNLVDLAGSER 240
D 177 ESHRHIGETNNVYSSRSHITFRMVIESREKVDSEAGESCD-AVRVSVLNLVDLAGSER 235
QY 241 AAGTGAAGVRLKEGCNINRSLFILGOVKKLSDG--QVGGFINVRDSKLTFRILQNSLGGN 298
D 236 AAKTGAAGVRLKEGSHINKSLMTLGTIVKKLSEGIEGQGHVYRDSKLTFRILQPALGGN 295
QY 299 PKTRIICTIT--PVSFDETLTALQFASTA 325
D 296 ANTALICNITLAQVHADETKSLQFASRA 324

RESULT 2
US-10-425-114-59725
; Sequence 59725, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53131)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59725
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3912-011-C4_FLI.pep
US-10-425-114-59725
Query Match 45.7%; Score 770.5; DB 12; Length 694;
Best Local Similarity 50.1%; Pred. No. 2e-70;
Matches 168; Conservative 53; Mismatches 97; Indels 17; Gaps 7;
QY 6 VAVCVVRP-----LNSREESLGETAQVYWKTDN--NVIIYQVDG--SKSFNDRVFEHG 54
D 30 ISVAVRFPNPLVAANTSPASSGGGDRWRIDTRVSLDLHRAAGPITGASFADHVFQ 89
QY 55 NETKNVYEEIAPITDSAIQYNGTIFAYGQTASGKTYTMMGSEDLHGVIPRAIHDIFQ 114
D 90 AANREIYGVVREILGAVGGFNGTAFAYGQTSKGITFMNGSDADPGIIPRAVRVFD 149
QY 115 KIKFPDRFLRVSYMEIYNETITDLGCTQKMKPLIIRVDNRRVYVADLTVEVVYVS 174
D 150 TVRQADREFLRVSYMEIYNEINDLL--TLGQKLTHESLDRGVYVSGUREIIVNSA 207
QY 175 ENALKWITKGEKSRHYGETKQNRSSRSHITFRMILSRKGPSSNCEGSVKVSHLNLVD 234
D 208 EQVFELLQGEANRHFGETNNRNRSSRSHITFRMVIESRGKQDTGGD-AIRVSVLNLVD 266
QY 235 LAGSRAAQTGAAGVRLKEGCNINRSLFILGOVKKLSD-QVGGFINVRDSKLTFRILQ 293
D 267 LAGSERIIKTGAEGVRLNEGKINKSLMILGNVINKLSENGKQGRGHIPYRDSKLTFR 326
QY 294 SLGGNPKTRIICTITP--VSFDETLTALQFASTAK 326
D 327 ALGNGAKTSIICTAAPEEIHIEETRGTLQFASRAK 361

RESULT 3
```

```
US-10-437-963-147933
; Sequence 147933, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 147933
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1006)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_48414C.1.pep
US-10-437-963-147933
Query Match 39.4%; Score 664; DB 16; Length 1006;
Best Local Similarity 45.8%; Pred. No. 3.9e-59;
Matches 153; Conservative 62; Mismatches 105; Indels 14; Gaps 9;
QY 1 AEGCAVAVCVVRPLNRSREESLGETAQVYWKTDNNVIIYQVDGSK-SFNFDREVHGNETTK 59
D 64 AAKESVAVAVRFPRLSPREVRGE--KIAWYADGETVARSEQSNLAYAYDRVFGPTTTR 121
QY 60 NYVEEIAPIITDSAIQYNGTIFAYGQTASGKTYTMMGSEDLHGVIPRAIHDIFQIKKF 119
D 122 HIYDAVAYVVGWAKNGINGTIFAYGVTSSKTHMGDQISPGVPLAVKDFINLIQET 181
QY 120 PDREFLLRVSYMEIYNETITDLGCTQKMKPLIIRVDNRRVYVADLTVEVVYVSEMALK 179
D 182 PNREFLLRVSYLEIYNEVNDLL--NPAGQNLRIREDLQGTI-VEGIKEEAVLSPVHALS 238
QY 180 WITKGE---KSRHYGETKQNRSSRSHITFRMILSRKGPSSNCEGSVKVSHLNLVDIA 236
D 239 LIAAGEVFSELRHVGSTNFIILSSRSHTIFITLTIESPRGQSNAE-AVTLSQLNLIDLA 297
QY 237 GSRAAQTGAAGVRLKEGCNINRSLFILGOVKKLSDQGVGGFINVRDSKLTFRILQNSIG 296
D 298 GSE--SSRVETAGVHQKEGSIYKSLTLGKVISKLT-D-EKATHIPFRDSKLTFRILKSSLS 355
QY 297 GNPKTRIICTITPVS--FDETLTALQFASTAKYM 328
D 356 GQGRVSLICTVTPASSNSEETHNLKFAHRAKHI 389

RESULT 4
US-10-437-963-182113
; Sequence 182113, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
```

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 182113
; LENGTH: 956
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(956)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_79331C.1.pep
US-10-437-963-182113

Query Match          38.6%; Score 651; DB 16; Length 956;
Best Local Similarity 46.6%; Pred. No. 8e-58;
Matches 156; Conservative 51; Mismatches 106; Indels 22; Gaps 8;

QY   2  BEGAVCVCRVRPLNGREESLGHTAQVYWKTDNN--VIYQVD-----GSKSFNFDRVPHG 54
Db   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      31 KEKIPVTVRVRPLSKKELAKD--QVAWECDNQTLIYGPPDPDRAAPTSYTFDKVFGP 88
QY   55 NETKNVYEIAAPIIDSAIQYNGTIFAYGQTASGKTYTMGSEDLHGVIPRAIHDFQ 114
Db   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      89 ASQTEVVYVEGAKDVAMSALTGINATIFAYGTSSGKTFTW-----GYTESAVNDIYR 142
QY   115 KIKKPPDRFELLRVSMELYNETITDLLCGTQKKPLIREDVNRNVYADLTEEVSYS 174
Db   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      143 HIENTPERDFIIKSAMISAMEYNEIVTKOLL--RPESNTLRLLDDPEKGTIVEKLEEEIAKDS 200
QY   175 EMALKWTIKGEKSRHYGETMKQRORSRSHTIFRMILESREREKGPSCEGSVKYSHMLVD 234
Db   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      201 QHLRLHSICESQRQVGETALNDTSRSQILRLTVESRLR-EVSGCVKSF-VANLNFDV 258
QY   235 LAGSRAAQTGAAGVRLKEGCNINRSLFIIGQVKKLSDGQVGFNYRDSKLTRILLONS 294
Db   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      259 LAGSRAAAQTHAVGARLKEGCHINRSULLTTVIRKLSDDRSGHIPYRDSKULTRIQLS 318
QY   295 LGGNPKTRLICTITPVSF---DETTLALOFASTAK 326
Db   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      319 LGGNARTAICTMSPAQTHVEQSRTXLFATCAK 353

RESULT 5
US-10-425-114-62672
; Sequence 62672, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53131)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62672
; LENGTH: 955
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-241-All_FLI pep
US-10-425-114-62672

Query Match          38.3%; Score 645.5; DB 12; Length 955;
Best Local Similarity 46.9%; Pred. No. 3e-57;

```

```

; SEQ ID NO 8
; LENGTH: 1235
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-143-8

Query Match      38.1%; Score 643; DB 15; Length 1235;
Best Local Similarity 44.4%; Pred. No. 8.e-57;
Matches 147; Conservative 51; Mismatches 117; Indels 16; Gaps 5;

QY 6 VAVCVVRPLNSREESLG-ETAQVYWKTDNNVYQVDSKSFNDFRVFHHGNETTKNVYEE 64
DB 13 VRVALRCRPLVPKEISEGCMCLSFVPGEPQVVGTG--KSFTYDFVDFPSTQEVEENT 70
QY 65 IAAPITDSAIQGYNGTIFAYGQTASGKTYTMMG-----SEHGLGVIPRAIHDFQKIK 117
DB 71 AVAPLKGVPKYNATVLAQGQSGKTYSMGAYTAEQENETPTGVIPRVQLLFKEID 130
QY 118 KFPDFRFLRVSYMEIYNETITDLCGQTKMKPLIIREDVNRNVYVADLTERRVYVTSMA 177
DB 131 KXSDFEFTLVSYLEIYNEEILDLLCPREKQINIREDPKEGKIIVGLTEKTVLVALDT 190
QY 178 LKWIITKGEKSRHYGETKNORSRSHITFRMILRESREKPEPNCESGVKVSHLNLDVLAG 237
DB 191 VSCLEQGNNSRVASTAMNSQSRSHAIPTISLEQRKKSD----KNSSFRSKLHLVDLAG 246
QY 238 SERAAQTGAAGVRLKEGNCINRSFILGOVVKKLSGQGVGFYINRDSKLTILQNSLGG 297
DB 247 SERQKTKAEGDRLKEGINNRGLCLGNVISALGDDKGGFVPYRDSKLTILQNSLGG 306
QY 298 NPKTRIITCTIPV--SFDETLTALQFASTAK 326
DB 307 NSHTLMIACVSPADSNLEETLNTLRYADRAR 337

RESULT 9
US-10-437-963-114373
; Sequence 114373, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 114373
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_18070C.1.pep
US-10-437-963-114373

Query Match      38.0%; Score 641.5; DB 16; Length 420;
Best Local Similarity 44.1%; Pred. No. 2.2e-57;
Matches 146; Conservative 56; Mismatches 108; Indels 21; Gaps 6;

QY 6 VAVCVVRPLNSREESLG-ETAQVYWKTDNNVYQVDSKSFNDFRVFHHGNETTKNVYEE 57
DB 17 IVVSVRLRPVNAREERGDSG--WECAGTTLTFRGAVPERAMFPASYSYDRVFSHECG 74
QY 58 TKNVYETIAPIIDSATQGYNGTIFAYGQTASGKTYTMMGSEHGLGVIPRAIHDFQKIK 117
DB 75 TRQYDEGARQVAMSVLSGINASTIFAYGQTSSGKTYTMMGITEY-----SMSDIYDYE 128

; SEQ ID NO 8
; LENGTH: 1235
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-143-8

Query Match      38.1%; Score 643; DB 16; Length 1232;
Best Local Similarity 44.4%; Pred. No. 8e-57;
Matches 147; Conservative 51; Mismatches 117; Indels 16; Gaps 5;

QY 6 VAVCVVRPLNSREESLG-ETAQVYWKTDNNVYQVDSKSFNDFRVFHHGNETTKNVYEE 64
DB 10 VRVALRCRPLVPKEISEGCMCLSFVPGEPQVVGTG--KSFTYDFVDFPSTQEVEENT 67
QY 65 IAAPITDSAIQGYNGTIFAYGQTASGKTYTMMG-----SEHGLGVIPRAIHDFQKIK 117
DB 68 AVAPLKGVPKYNATVLAQGQSGKTYSMGAYTAEQENETPTGVIPRVQLLFKEID 127
QY 118 KFPDFRFLRVSYMEIYNETITDLCGQTKMKPLIIREDVNRNVYVADLTERRVYVTSMA 177
DB 128 KXSDFEFTLVSYLEIYNEEILDLLCPREKQINIREDPKEGKIIVGLTEKTVLVALDT 187
QY 178 LKWIITKGEKSRHYGETKNORSRSHITFRMILRESREKPEPNCESGVKVSHLNLDVLAG 237
DB 188 VSCLEQGNNSRVASTAMNSQSRSHAIPTISLEQRKKSD----KNSSFRSKLHLVDLAG 243
QY 238 SERAAQTGAAGVRLKEGNCINRSFILGOVVKKLSGQGVGFYINRDSKLTILQNSLGG 297
DB 244 SERQKTKAEGDRLKEGINNRGLCLGNVISALGDDKGGFVPYRDSKLTILQNSLGG 303
QY 298 NPKTRIITCTIPV--SFDETLTALQFASTAK 326
DB 304 NSHTLMIACVSPADSNLEETLNTLRYADRAR 334

RESULT 8
US-10-334-143-8
; Sequence 8, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; CURRENT FILING DATE: 2002-12-31
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1

```





Db 296 LQDSLGGNTKILMVACLSPADNNYDETSLRYANRAK 333

RESULT 14

US-10-311-642-4

Sequence 4, Application US/10311642

Publication No. US20040086878A1

GENERAL INFORMATION:

APPLICANT: Cytokinetics, Inc.

APPLICANT: Beraud, Christophe

APPLICANT: Freedman, Richard

TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR

FILE REFERENCE: 020552-001910US

CURRENT APPLICATION NUMBER: US/10/311,642

CURRENT FILING DATE: 2003-09-29

PRIOR APPLICATION NUMBER: 09/597,602

PRIOR FILING DATE: 2000-06-20

NUMBER OF SEQ ID NOS: 4

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 4

LENGTH: 329

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Amino acid sequence of preferred fragment of motor

OTHER INFORMATION: domain of Hskif17

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Hskif17

US-10-311-642-4

Query Match 36.8%; Score 620.5; DB 16; Length 329;

Best Local Similarity 42.7%; Pred. No. 2.2e-55;

Matches 144; Conservative 54; Mismatches 114; Indels 25; Gaps 8;

Qy 6 VAVCVVRPLNSREESLGETAQVYWKTD-----NNVIYQVDSKSFNDRVPHGNET 57

Db 1 KVIVVRCRPMQREEL--RCQPVTVVDCARAQCCIQNPGAADPPKQFTFDGAYHVDH 58

Qy 58 TKNVVEEIAPIIDSAIQYNGTIFAYGQTASGKTYTMGSED---HLGVIPRAIHDF 114

Db 59 TEQIYNEIAYPLVEGVTEGYNGTIFAYGQTSGKSFYTMQGLPDPSPQRGIIPRAFEHVF 118

Qy 115 KIKKPPDFRFLRVSYMEIYNETITDLCGTCQKMKPLIREDVNNVYVADLTSEVWYT 174

Db 119 SVQCAENTKFLVRASYLEIYNEVDRLDGADTKQK-LELKEHPEKGVYVKGLSMHTVHSV 177

Qy 175 EMALKWITKGEKSRHVGTEKMNORSRSHTIFRMILESR---EKGEPSNCEGSVKVSHLN 231

Db 178 AQCEHIMETGWNKRVGVYTLNMKDSRSRHSITISIEMSAVDERG-----KDHLAGKLN 232

Qy 232 LVDLAGSRAAQTGAAGVRLKEGCNINRSLFILGVIKLSQGVGGFYNIRDSKLTRL 291

Db 233 LVDLAGSERQSKTGATGERLKEATKINLSLALGNVISALVDGRC-KHVPYRDSKLTLL 291

Qy 292 QNSLGNPKTRIICITIPV--SFDETLTALQFASTAK 326

Db 292 QDSLGGNTKILMVACLSPADNNYDETSLRYANRAK 328

RESULT 15

US-10-437-963-169150

Sequence 169150, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

Qy 65 IAAPIDSALQYNGTIFAYGQTASGKTYTMG-----SEHGLVIPRAIHDFQKIK 117

Db 69 AVAPLIKGVFKGNATVLAAYGQSGKTYSMGAYTAEGENPTGVIPRVQLLPKEID 127

Qy 118 KFPDFRFLRVSYMEIYNETITDLCGTCQKMKPLIREDVNNVYVADLTSEVWYTSEMA 177

Db 128 KKSDFEFTLKVSLEYINEBEILLCPREKAQINIREDPKEGIKIVGLTEKTVLVALDT 187

Qy 178 LKWITKGEKSRHVGTEKMNORSRSHTIFRMILESRKESKPSNCEGSVKVSHNLVDLAG 237

Db 188 VSLQEQNNRSTVASTAMNSQSRSHAFTI---SLEQKSKDNKSSFR-SKLHLVDLAG 243

Qy 238 SERAAQTGAAGVRLKEGCNINRSLFILGVIKLSQGVGGFYNIRDSKLTRLQNSLGG 297

Db 244 SEQOKTKAEGDLKGININRGLCLGNVISALGDDKGGFAPYRDSKLTRLLODSLGG 303

Qy 298 NPKTRIICITIPV--SFDETLTALQFASTAK 326

Db 304 NSHTLMIACVSPADSNLEETINTLRYADRAR 334

RESULT 13

US-10-311-642-2

Sequence 2, Application US/10311642

Publication No. US20040086878A1

GENERAL INFORMATION:

APPLICANT: Cytokinetics, Inc.

APPLICANT: Beraud, Christophe

APPLICANT: Freedman, Richard

TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR

FILE REFERENCE: 020552-001910US

CURRENT APPLICATION NUMBER: US/10/311,642

CURRENT FILING DATE: 2003-09-29

PRIOR APPLICATION NUMBER: 09/597,602

PRIOR FILING DATE: 2000-06-20

NUMBER OF SEQ ID NOS: 4

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 2

LENGTH: 1029

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Hskif17 amino acid sequence

OTHER INFORMATION: Description of Artificial Sequence: Hskif17

US-10-311-642-2

Query Match 37.0%; Score 624.5; DB 16; Length 1029;

Best Local Similarity 42.9%; Pred. No. 5e-55;

Matches 145; Conservative 54; Mismatches 114; Indels 25; Gaps 8;

Qy 5 AVAVCVVRPLNSREESLGETAQVYWKTD-----NNVIYQVDSKSFNDRVPHGNE 56

Db 5 AVKVVRCRPMQREEL--RCQPVTVVDCARAQCCIQNPGAADPPKQFTFDGAYHVDH 62

Qy 57 TKNVVEEIAPIIDSAIQYNGTIFAYGQTASGKTYTMGSED---HLGVIPRAIHDF 113

Db 63 VTEQIYNEIAYPLVEGVTEGYNGTIFAYGQTSGKSFYTMQGLPDPSPQRGIIPRAFEHVF 122

Qy 114 QIKKPPDFRFLRVSYMEIYNETITDLCGTCQKMKPLIREDVNNVYVADLTSEVWYT 173

Db 123 ESQCAENTKFLVRASYLEIYNEVDRLDGADTKQK-LELKEHPEKGVYVKGLSMHTVHS 181

Qy 174 SEMALKWITKGEKSRHVGTEKMNORSRSHTIFRMILESR---EKGEPSNCEGSVKVSHL 230

Db 182 VAQCEHIMETGWNKRVGVYTLNMKDSRSRHSITISIEMSAVDERG-----KDHLAGKLN 236

Qy 231 NLVDLAGSRAAQTGAAGVRLKEGCNINRSLFILGVIKLSQGVGGFYNIRDSKLTTRI 290

Db 237 NLVDLAGSERQSKTGATGERLKEATKINLSLALGNVISALVDGRC-KHVPYRDSKLTLL 295

Qy 291 LQNSLGNPKTRIICITIPV--SFDETLTALQFASTAK 326





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 09:28:21 ; Search time 4.90075 Seconds

(without alignments)

6437.961 Million cell updates/sec

Title: US-10-045-631B-88\_COPY\_2\_329

Perfect score: 1686

Sequence: 1 AEGGAVAVCVVRPLNSREE.....PVSFDETLTALQFASTAKYM 328

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1686	100.0	2663	1	S28261
2	1258.5	74.6	2954	2	T14156
3	773.5	45.9	823	2	T52425
4	738	43.8	888	2	D96619
5	663.5	39.4	1459	2	T30196
6	650	38.6	786	2	A53939
7	643	38.1	701	1	B44259
8	642	38.1	742	1	S58691
9	641.5	38.0	1231	2	A54803
10	630.5	37.4	909	2	H86350
11	630	37.4	747	1	A57107
12	628.5	37.3	1226	2	I51617
13	623.5	37.0	699	1	S38982
14	621	36.8	932	2	T49235
15	618.5	36.7	1121	2	T06065
16	612.5	36.3	968	2	T45746
17	608	36.1	1225	2	A56514
18	607	36.0	581	2	F84599
19	605	35.9	672	2	S54351
20	601.5	35.7	1263	2	T13465
21	597	35.4	332	2	C48835
22	595	35.3	1921	1	T13827
23	593.5	35.2	975	1	A31497
24	591.5	35.1	1031	1	A38713
25	591.5	35.1	1584	1	JN0114
26	591.5	35.1	1584	2	T15822
27	586.5	34.8	554	2	T50118
28	586	34.8	1254	2	T18277
29	585	34.7	793	2	JC5831

30	585	34.7	1076	2	B84687
31	584.5	34.7	1130	2	T21134
32	582	34.5	928	2	T10164
33	582	34.5	963	1	A41919
34	582	34.5	1066	1	A48669
35	580.5	34.4	1032	2	I38510
36	579.5	34.4	1388	2	T30335
37	579	34.3	968	2	T51933
38	578.5	34.3	294	2	S38983
39	576.5	34.2	967	1	A35075
40	573.5	34.0	1576	2	T29237
41	573	34.0	1027	2	S37711
42	572.5	34.0	834	2	T06055
43	571.5	33.9	935	2	T51930
44	569	33.7	843	2	S44868
45	568	33.7	784	1	A55236

probable kinesin-1  
hypothetical prote  
kinesin heavy chai  
kinesin heavy chai  
kinesin-related pr  
neural kinesin h  
KLP2 protein - Afr  
kinesin motor prot  
kinesin-related pr  
kinesin heavy chai  
hypothetical prote  
hypothetical prote  
kinesin (imported)  
kinesin heavy chai  
kinesin-related pr

## ALIGNMENTS

### RESULT 1

S28261

Centromere protein E - human

N;Alternate names: Centromere 312K protein; kinesin-related protein CENP-E

C;Species: Homo sapiens (man)

C;Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 19-Jan-2001

C;Accession: S28261

R;Yen, T.J.; Li, G.; Schaar, B.T.; Szilak, I.; Cleveland, D.W.

Nature 359, 536-539, 1992

A;Title: CENP-E is a putative kinetochore motor that accumulates just before mitosis.

A;Reference number: S28261; MUID:93024922; PMID:1406971

A;Accession: S28261

A;Molecule type: mRNA

A;Residues: 1-2663 <YEN>

A;Cross-references: EMBL:Z15005; NID:G29864; PIDN:CAA78727.1; PID:G29865

C;Genetics:

A;Gene: GDB:CENPE

A;Cross-references: GDB:361164; OMIM:117143

A;Map position: 4q24-4q25

C;Superfamily: centromere protein E; kinesin motor domain homology

C;Keywords: ATP; coiled coil; microtubule binding; mitosis; nucleotide binding; P-loop

F;7-335/Domain: kinesin motor domain homology <KNOT>

F;86-93/Region: nucleotide-binding motif A (P-loop)

F;486-2183/Domain: coiled coil #status predicted <COI>

F;92/Binding site: ATP (Lys) #status predicted

Query Match	100.0%	Score 1686;	DB 1;	Length 2663;
Best Local Similarity	100.0%;	Pred. No. 3.1e-120;	Mismatches 0;	Indels 0;
Matches 328;	Conservative 0;	Indels 0;	Gaps 0;	

QY	1	AEGGAVAVCVVRPLNSREEISLGETAQVYKTDNNVIYQVDSKSFENFDRVPHGNETYKN	60
Db	2	AEGGAVAVCVVRPLNSREEISLGETAQVYKTDNNVIYQVDSKSFENFDRVPHGNETYKN	61
QY	61	YVEHIAPIIDSAIQNGNTTIFAYGQTASGKTYTMGSEDLGVIPRAIHDFQKIKKFP	120
Db	62	YVEHIAPIIDSAIQNGNTTIFAYGQTASGKTYTMGSEDLGVIPRAIHDFQKIKKFP	121
QY	121	DRELLRVSYMEIYNETITDLGCTGCKMPLIREDVNRNVIYADLTVEVVYTTSEMAIKW	180
Db	122	DRELLRVSYMEIYNETITDLGCTGCKMPLIREDVNRNVIYADLTVEVVYTTSEMAIKW	181
QY	181	ITKGEKSRHYGETKMNQSSRSHTIFRMILESREKGPSPCEGSKVSHLNLDLAGSER	240
Db	182	ITKGEKSRHYGETKMNQSSRSHTIFRMILESREKGPSPCEGSKVSHLNLDLAGSER	241
QY	241	AAQTGAAGVRLKGGCNINRSLFILGQVIKSLDGQVGGFINYRDSKLTRILQNSLGNPK	300
Db	242	AAQTGAAGVRLKGGCNINRSLFILGQVIKSLDGQVGGFINYRDSKLTRILQNSLGNPK	301
QY	301	TRIICITTPVSFDETLTALQFASTAKYM	328



Db	236	VRFSRYERDLLLVITSR--GKDSSSDAIRVSVLNLVDLAGSERIAKTGAGVRVLRQEGKY	293
Qy	257	INRSLFILGVQIKKLSDG-QVGGFNYRDSKLTILQNSLGNPKTRICTITPVS--FD	313
Db	294	INSLMILGVNINKLSDTKLRAHPYRDSKLTILQPALGNNAKTCIICTIAPEEHIE	353
Qy	314	ETLTALQFASAK	326
Db	354	ESKGTLOFASRAK	366

RESULT 5  
T30196  
kinesin motor protein 1 - smut fungus (Ustilago maydis)  
C:Species: Ustilago maydis (corn smut)  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 11-May-2000  
C:Accession: T30196  
R:Lehmlier, C.; Steinberg, G.; Snetselaar, K.M.; Schliwa, M.; Kahmann, R.; Bolker, M.  
EMBL J. 16, 3464-3473, 1997  
A:Title: Identification of a motor protein required for filamentous growth in Ustilago  
A:Reference number: Z20770; MUID:97361828; PMID:9218789  
A:Accession: T30196  
A>Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-1459 <LEH>  
A:Cross-references: EMBL:U92844; NID:92062749; PID:92062750; PIDN:AA633336.1  
C:Genetics:

A;Description: required for filamentous growth in Ustilago maydis

Query Match	39.4%;	Score 663.5;	DB 2;	Length 1459;
Best Local Similarity	39.8%;	Pred. No. 2.2e-42;		
Matches 159;	Conservative	51;	Mismatches 100;	Indels 89; Gaps 10;

  

QY	6	VAVCVVRPLNSREESLGETAQVYWKTDN--NVIYQVD-----	41
Db	237	VVVCVRMP--SRASSDSEASV-WNCDSSEKRIFFTPEHHPAIAKRTTSSERAGAGASIA	293
QY	42	-----GSKSNFDRVRFHGNETTNNVBEIAAPIIDSAIQGVNGTIFAYGQ	86
Db	294	AAPSSHDLDHEDPTSSTYHFQPKLTIGAQTTDDMVHSHIAPVRAAVEGVNGVFAYGQ	353
QY	87	TASGKTYTWMGSEDLGVIPRAHDIFOKIKKPDREFLLRVSYMEIYNFTITDLLCGTG	146
Db	354	TSGKHTTWGSGDAEPGVIPRAVEQIFQMIKDEPDREFLLRVSYLEIYNETLKDLA---	410
QY	147	KMKPL-----IREDVNRNVVYADLTAEV	171
Db	411	PLPLPTGTSGSLQTTDRPASPIKGGSSHAAGSQSCTLRITEDQKSSRVITGLREEIV	470
QY	172	YTSSEALKWITGKSRHHGETKMQORSRSHTIFRMILESREKGBFSCGVSKVSHLN	231
Db	471	TDANTVLCIQGQDERHVGATDWNERSRSHCVFOLTIESRPPASASKE--VRISQLN	528
QY	232	LVDLAGSRAAQTAGAGVRLKEGCNINRSLFILGOVIKKLSDQVGG--FINVRDSKLTR	289
Db	529	LIDLAGSRAA---SQAERREKGAFLINKSLLTGTVIGKLTPEVNGDAHIPYRDSKLTR	585
QY	290	ILONSLGNGPKTRITCTTPVS--FDETLTALQFASAK	326
Db	586	ILQTSLGNARIAVICTISPDTEHANETITSLKFGKRCK	624

RESULT 6  
A53939  
kinesin homolog KHPI - Chlamydomonas reinhardtii  
C:Species: Chlamydomonas reinhardtii  
C/Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 02-Feb-2001  
C/Accession: A53939  
R/Walther, Z.; Vashisinha, M.; Hall, J.L.  
J. Cell Biol. 126, 175-188, 1994  
A/Title: The Chlamydomonas FLA10 gene encodes a novel kinesin-homologous protein.

A/Reference number: A53939; MUID:94299638; PMID:8027176  
A/Accession: A53939  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-786 <WAL>  
A/Cross-references: EMBL:I33697; NID:G497696; PIDN:AAA21738.1; PID:G497697  
A/Note: authors translated the codon AAC for residue 753 as Asp  
C/Genetics:  
A/Gene: FLA10  
C/Superfamily: unassigned kinesin-related proteins; kinesin motor domain  
C/Keywords: ATP; coiled coil; nucleotide binding; P-loop  
F/1-359/Domain: kinesin motor domain homology <KW07>  
F/97-104/Region: nucleotide-binding motif A (P-loop)

Query Match	38.6%;	Score 650;	DB 2;	Length 786;	
Best Local Similarity	44.8%;	Pred. No. 1e-41;			
Matches 154;	Conservative 56;	Mismatches 110;	Indels 24;	Gaps 8;	
Qy	5	AVAVCVRVRLPNSREESLGETAQVYWKTDNNVY----	QVDGS----	KSFNPRDRVHGNET 57	
Db	10	SVKVVVRCRLNGKEKADGHSRIVDMVDVDAQVKVRNPKADASEPPKAF	T	FDQVYDMNCQ 69	
Qy	58	TKNVEIEAIIIDSATQGYNGTIFAYGQTASGKTYTMMGSEDH----	LGVI	PRAIHDIFQ 114	
Db	70	QRDVFDIRPLDSCIEGNGTIFAYGQTGOKSHMEGKDPPELGRULIPN	FRYVFE	129	
Qy	115	KI-KKFPDRFLRVSMYMEIYNTITDLCGTOKMKPLIIREDVNRN	VYVADL	TEEVVYT 173	
Db	130	IIARDSGTKEFLVRSSYLEIYNEEVDLL-KGDH	SKKMKELKES	PDGRGVVYKDL	LSQFVCKN 188
Qy	174	SEMAKWIWTKGSRHYGETKNNORSRSHITPRMILESEK-----	GP	SNCGECS 224	
Db	189	YEEMKVLVLAGKNRVQVGTALMMQDSRSRHSIPTITIECIEKLE	S	AAAKPGAKKDDSNH 248	
Qy	225	VKVSHLNLVDLAGSEAAQTGAAGVRLKEGCNINRSLFTILGOV	IKKLS	DGVGVGFNYRD 284	
Db	249	VRVGKLNLDLAGSERQDKTIGATGRDLKESGKINLSLTALGNV	ISALVQK-SG	HIPIYRD 307	
Qy	285	SKLTRLQNSLGNPKTRIICITPV--SFDETLTALQFASTAK		326	
Db	308	SKLTRLQDSLGNTKTVMANIGADWNYDETMSTLRYANRAK		351	

```

RESULT 7
B44259
kinasin-related protein KIF3A - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 19-Jan-2001
C:Accession: B44259; S27872
R:Aizawa, H.; Sekine, Y.; Takemura, R.; Zhang, Z.; Nangaku, M.; Hirokawa, N.
J. Cell Biol. 119, 1287-1296, 1992
A:Title: Kinesin family in murine central nervous system.
A:Reference number: A44259; MUID:93077686; PMID:1447303
A:Accession: B44259
A:Molecule type: mRNA
A:Residues: 1-701 <AIZ>
A:Cross-references: EMBL:D12645; NID:G220469; PIDD:BA02166.1; PID:G220470
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:118911)
C:Complex: heterodimer with KIF3B (PIR:A57107); the KIF3A/3B heterodimer ass
C:Function:
A:Description: KIF3 complex is a motor protein that provides anterograde fac
C:Superfamily: kinesin-related motor KIF3; kinesin motor domain homology
C:Keywords: ATP; coiled coil; heterodimer; heterotrimer; microtubule binding
F:1-368/Domain: head globular #status predicted <HGL>
F:15-351/Domain: kinesin motor domain homology <KNOT>
F:100-107/Region: nucleotide-binding motif A (p-loop)
F:369-599/Domain: helical rod #status predicted <ROD>
F:600-701/Domain: tail globular #status predicted <TGL>
F:106/Binding site: ATP (Iys) #status predicted

```

Query Match 38.1%; Score 643; DB 1; Length 701;  
Best Local Similarity 46.2%; Pred. No. 3.1e-41;

Matches 154; Conservative 48; Mismatches 115; Indels 16; Gaps 8;  
QY 6 VAVCVVRPLNREESL--GETAQVYWKTDNNVIYQVDSGSGS---KSFNFDVRFHGNFTTK 59  
DB 15 VKVWVRCRPLNREKESNCVRAQVSDVEMRGITIVHKTDSSNEPPKFTFDVTFVGPESKQL 74  
QY 60 NVYEELAAPIDSAIOGYNGTIFAYGQTASGKTYTWGSE---DHGVIPRAIHDFOKI 116  
DB 75 DVYNLTARPLIDSVLEGYNGTIFAYGQTGKTFTMEGVRAVPLGRGVIPNSFAHIFGHI 134  
QY 117 KKPP-DREFLLRVSYMEIYNETITDILCGTQKMKPLIIRDVNRNYYVADLTBEVVYTS 175  
DB 135 AKAEGRDTRFVRVSYLIEYNEERDGL-GDQQRLEVRKRPDVGVTIKDLSAYVNNAD 193  
QY 176 MALKWITKGEKSHYGETKMNORSSRSHITFRMILESRKGEPSNCEGSKVSHLNLVDL 235  
DB 194 DMDRIMTLGHKNRSGVATNNHSSRSHAFITIECESEKGVGDGMH--VRMGKHLVDL 251  
QY 236 AGSERRAQTGAAGVRLKEGNCINRSLFILGQVTKKLSGQGVGFPIFYRDSKLTLLONSL 295  
DB 252 AGSERQAKTGATGQRLKEATKINLSLSTLGNVISALVDGK-STHVPYRNSKLTLLQDSL 310  
QY 296 GGNPKTRIICTITPV--SPDETTLAQFASTAK 326  
DB 311 GGNSTWMCANIGPADNYDETITLRYANRAK 343

## RESULT 8

S58691  
A: kinein-related protein KRP95 - sea urchin (Strongylocentrotus droebachiensis)  
N: Alternate names: kinesin-2 chain B; KRP (85/95) 95K chain  
C: Species: Strongylocentrotus droebachiensis  
C: Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 19-Jan-2001  
C: Accession: S58691  
J: Rashid, D.J.; Wedaman, K.P.; Scholey, J.M.  
J. Mol. Biol. 252, 157-162, 1995  
A: Title: Heterodimerization of the two motor subunits of the heterotrimeric kinesin, KRP95  
A: Reference number: S58691; MUID: 95404610; PMID: 7674298  
A: Accession: S58691  
A: Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation  
A: Molecule type: mRNA  
A: Residues: 1-742 <RAS>  
C: Complex: heterotrimer of a 115K chain and two kinesin-related chains of 85K (PIR:S3898)  
C: Superfamily: kinesin-related protein KIF3; kinesin motor domain homology  
C: Keywords: ATP; heterotrimer; microtubule binding; nucleotide binding; P-loop  
F: 9-345/Domain: kinesin motor domain homology <KMOT>  
F: 95-102/Region: nucleotide-binding motif A (P-loop)  
F: 101/Binding site: ATP (lys) #status predicted

Query Match 38.1%; Score 642; DB 1; Length 742;  
Best Local Similarity 45.0%; Pred. No. 3.9e-41;  
Matches 150; Conservative 50; Mismatches 117; Indels 16; Gaps 6;  
QY 6 VAVCVVRPLNREESLGETAQVYWKTDNNVIYQV-----DGSKSFNFDVRFHGNFTT 58  
DB 9 VKVWVRCRPMNSKEISQGHKRVEMDNKGLVEVTNPKGPPGPKNSFTFDIVYDWSKQ 68  
QY 59 KNVYEELAAPIDSAIOGYNGTIFAYGQTASGKTYTWG---SEDLGVIPRAIHDFOKI 115  
DB 69 IDLYDTFRSLVESVLOGNGTIFAYGQTGKTFTMEGVRSNPRLRGVIPSFEHIFTH 128  
QY 116 IKFPDPREFLLRVSYMEIYNETITDILCGTQKMKPLIIRDVNRNYYVADLTBEVVYTS 175  
DB 129 IARTNQOQLVRASYLIEYQBEIRDLAKDQK-KRLDLKERPDVTGYVVDLSSFTVTSVK 187  
QY 176 MALKWITKGEKSHYGETKMNORSSRSHITFRMILESRKGEPSNCEGSKVSHLNLVDL 235  
DB 188 EIEHVMVTGNRNSVSGTNNHSSRSHAFITIECESELGVDG--ENHIRVGKJNLVDL 245  
QY 236 AGSERRAQTGAAGVRLKEGNCINRSLFILGQVTKKLSGQGVGFPIFYRDSKLTLLONSL 295  
DB 246 AGSERQAKTGATGRLKEATKINLSLSTLGNVISALVDGK-SSHIPYRDSKLTLLQDSL 304

QY 296 GGNPKTRIICTITPV--SPDETTLAQFASTAK 326  
DB 305 GGNAKTVWANNGPASYNFEDEITITLRYANRAK 337  
RESULT 9  
A54803  
N: microtubule-associated motor KIF4 - mouse  
C: Alternate names: kinesin-related protein KIF4  
C: Species: Mus musculus (house mouse)  
C: Date: 05-Apr-1995 #sequence\_revision 05-Apr-1995 #text\_change 02-Feb-2001  
C: Accession: A54803; D44259  
R: Sekine, Y.; Okada, Y.; Noda, Y.; Kondo, S.; Aizawa, H.; Takemura, R.; Hirokawa, N.  
J. Cell Biol. 127, 187-201, 1994  
A: Title: A novel microtubule-based motor protein (KIF4) for organelle transports, whose  
A: Reference number: A54803; MUID: 95014709; PMID: 7929562  
A: Accession: A54803  
A: Molecule type: mRNA  
A: Residues: 1-1231 <SEK>  
A: Cross-references: GB: D12646; NID: G563772; PIDN: BAA02167.1; PID: d1002657; PID: G563773  
R: Aizawa, H.; Sekine, Y.; Takemura, R.; Zhang, Z.; Nangaku, M.; Hirokawa, N.  
J. Cell Biol. 119, 1287-1296, 1992  
A: Title: Kinesin family in murine central nervous system.  
A: Reference number: A44259; MUID: 93077686; PMID: 1447303  
A: Accession: D44259  
A: Status: nucleic acid sequence not shown; not compared with conceptual translation  
A: Molecule type: mRNA  
A: Residues: 91-111, 'S', 113-240 <AIZ>  
A: Experimental source: brain  
A: Note: sequence extracted from NCBI backbone (NCBIP: 118904)  
C: Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology  
C: Keywords: ATP; nucleotide binding; P-loop  
F: 10-343/Domain: kinesin motor domain homology <KMOT>  
F: 88-95/Region: nucleotide-binding motif A (P-loop)

Query Match 38.0%; Score 641.5; DB 2; Length 1231;  
Best Local Similarity 44.3%; Pred. No. 8.4e-41;  
Matches 147; Conservative 53; Mismatches 115; Indels 17; Gaps 7;  
QY 6 VAVCVVRPLNREESLQ-ETAQVYWKTDNNVIYQVDSGSKSFNFDVRFHGNFTTKNVYEE 64  
DB 10 VRVALRCKPLVSKIEKGCQCLSFVGPQVQV--VGNDKSTFYDFVDPSTQEVEVT 67  
QY 65 IAAPIIDSAIOGYNGTIFAYGQTASGKTYTWGSGS---EDH---LGVIPRAIHDFOKIK 117  
DB 68 AVAPLKGFGYNATVLAYGQTGSGKTYSMGGAYTAQEHDSATGIVPRVQLLFKEIN 127  
QY 118 KPDPREFLLRVSYMEIYNETITDILCGT-QKMKPLIIRDVNRNYYVADLTBEVVYTS 176  
DB 128 KKSDFEFTLRVSYLIEYNEEILDLCSREKATQINREDPKEGKIYGLIEKTVLVASD 187  
QY 177 ALKWTIKGEKSHYGETKMNORSSRSHITFRMILESRKGEPSNCEGSKVSHLNLVDLA 236  
DB 188 TVSCLEQQGNERTVASTAMNSQSSRSHAFITISIQKXK---NDKNSFRSKLHLVDLA 243  
QY 237 GSERAAQTGAAGVRLKEGNCINRSLFILGQVTKKLSGQGVGFPIFYRDSKLTLLONSL 296  
DB 244 GSERQKTKAEGRDLREGININRGLLCLGNVISALGDDKKGNFVYRDSKLTLLQDSL 303  
QY 297 GNPKTRIICTITPV--SPDETTLAQFASTAK 326  
DB 304 GNSHTIMTACVSPADSNLEETLTLRYADRAR 335

## RESULT 10

H86350  
N: hypothetical protein F9K7.17 - Arabidopsis thaliana  
C: Species: Arabidopsis thaliana (mouse-ear cress)  
C: Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
C: Accession: H86350  
R: Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.  
ansen, N.F.; Hughes, B.; Huizar, L.



<b>Query Match</b>							
37.4%; Score 630; DB 1; Length 747;							
<b>Best Local Similarity</b> 43.8%; Pred. No. 3.3e-40;							
<b>Matches</b> 147; Conservative 54; Mismatches 115; Indels 20; Gaps 7;							
QY	5	AVAVCVVRPLNSREESLG-----ETAQVYWKTDNNVIYQVDGSKSFNFDRVFHGN	55				
Dd	9	SVRVVRCRPMNGEKAASVDKVVDKLGVQVKPKGTSHBM--PKTFTEDAVDWN	66				
QY	56	ETTKNVYEIAAPLIDSAIGYNGTIFAYQTASGKTYYM---MGSEDLGLGVIPRAIHD	112				
Dd	67	AKOPELYDETFRPLDVDVLQFNGTIFAYQGTGKTYYMEGVGRDPEKGIVPNSFDHI	126				
QY	113	FQIKKKPPDEFLLRVSYMEIYNETHDLLCGTGQKWKLPIREDVNARNVYVADLTBEVVY	172				
Dd	127	FTHLRSQNQYLVRASYLEYIOEERIDLDSKDQ-TKRLELKRPDTGGVYVKDLSFPVTK	185				
QY	173	TSEMALKWITFKBKSRYHTGTKNORSSRSHTLFRMLIESREKGEFSNCEGSVKVSHLN	232				
Dd	186	SVKEIEHVNMVNGNSVGATNWNEHSSRSHALFVITIECSVG--LDGENHIRVGKLN	243				
QY	233	VDLAGSERAAQTGAAGVRUKEGCNINRSFLIGQVTKLSDGQGVGFNVNRDSKLTRLIQ	292				
Dd	244	VDLAGSEROAKTGAOGERLKEATKINLSISALGNVISALVDGK-STHIPIYRDSKLTRLIQ	302				
QY	293	NSLGGNPKTRIICITTPVSF--DETLTAQOPASTAK	326				
Dd	303	DSLGGNAKVWVANVPASNVEETLTILRYANRAK	338				
<b>RESULT 12</b>							
I51617							
kinesin-like protein 1 - African clawed frog							
C:Species: Xenopus laevis (African clawed frog)							
C>Date: 13-Sep-1996 #sequence revision 13-Sep-1996 #text_change 02-Feb-2001							
C:Accession: I51617; A48835; S48837							
R:Vernos, I.; Raats, J.; Hiranou, T.; Heasman, J.; Karsenti, E.; Wylie, C.							
Cell 81, 117-127, 1995							
A>Title: Xklp1, a chromosomal Xenopus kinesin-like protein essential for spindle							
A:Reference number: A56221; PMID:95236444; PMID:7720067							
A:Accession: I51617							
A>Status: preliminary; translated from GB/EMBL/DDBJ							
A:Molecule type: mRNA							
A:Residues: 1-1226 <VER>							
A:CROSS-references: EMBL:X82012; NID:G562792; PIDN:CAA57539.1; PTD:G562793							
Dev. Biol. 157, 232-239, 1993							
A>Title: Multiple kinesin-like transcripts in Xenopus oocytes.							
A:Reference number: A48835; MUID:93246065; PMID:8482413							
A:Accession: A48835							
A>Status: preliminary; not compared with conceptual translation							
A:Molecule type: mRNA							
A:Residues: 9-162 'L', 164-338 <VE2>							
A:Experimental source: oocyte							
A>Note: sequence extracted from NCBI backbone (NCBIP:I30975)							
C:Genetics:							
A:Gene: klp1							
A:Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology							
C:Keywords: ATP; nucleotide binding; p-loop							
C:F:9-343/Domain: kinesin motor domain homology <KMOT>							
C:F:87-94/Region: nucleotide-binding motif A (p-loop)							
<b>Query Match</b>							
37.3%; Score 628.5; DB 2; Length 1226;							
<b>Best Local Similarity</b> 44.1%; Pred. No. 8.3e-40;							
<b>Matches</b> 149; Conservative 53; Mismatches 117; Indels 19; Gaps 9;							
QY	2	BEG-AVAVCVVRPLNSRESLG-ETAQVYWKTDNNVIYQVDGSKSFNFDRVFHGN	59				
Dd	4	DEGIPIRVVALRCRPLVPKENNECKMCLTFVPCQQVI-VGTEKFTDYDFPSAEQE	61				
QY	60	NVYEETAAPLIDSAIGYNGTIFAYQTASGKTYYTMGSEDH-----LGVIPRAIHDI	112				
Dd	62	EVDNSAVAPLIKGLFGKYNAVTLVAYQSGTKYSMGAYTHNQENPTVGVIPRTVAL	121				

113 FOKIKKPPDEFLLRVSYMEIYNETITDLL-CGTQKMKFLIREDVNRNVYADLTVEV 171  
 122 FREIHQRPWFNKLVSLEYINEELDLLYAARDKNTISIREDPKEGKICGLTERDV 181  
 172 YTSSEALKWITTKGSKSRHYGETKMNORSRSHITFRMILESKRGPSCEGSKVYSHLN 231  
 182 KTAIDTLSCLEQNSRSTVASTAMNSQSSRSHAFITISIEQRKEGKNN---SFR-SKLH 237  
 232 LVDLAGSRAAQTGAAGVRLKEGCNINRSFILGQVKKLSD-GQVGGFNYRDSKLTRI 290  
 238 LVDLAGSERQKTKAEDRKEGISNRGLCLGNVISALGDSKGGFVYRDSKLTRL 297  
 291 LQNSLGNPKTRIICTTPV--SFDETLTALQFASAK 326  
 298 LQNSLGNSHTLMAICVSPADSNMEETLNTLRVADRAR 335

RESULT 13  
 S38982  
 kinesin-related protein KRP85 - sea urchin (Strongylocentrotus purpuratus)  
 N:Alternate names: kinesin-2 chain A; KRP (85/95) 85K chain  
 C:Species: Strongylocentrotus purpuratus (purple urchin)  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 19-Jan-2001  
 C:Accession: S38982; S72551  
 R:Cole, D.G.; Chinn, S.W.; Wedaman, K.P.; Hall, K.; Vuong, T.; Scholey, J.M.  
 Nature 366, 268-270, 1993  
 A:Title: Novel heterotrimeric kinesin-related protein purified from sea urchin eggs.  
 A:Reference number: S38982; MUID:94050179; PMID:8232586  
 A:Accession: S38982  
 A:Molecule type: mRNA  
 A:Residues: 1-699 <COL1>  
 A:Cross-references: EMBL:L16993; NID:g295245; PIDN:AAA16098.1; PID:g295246  
 A:Accession: S72551  
 A:Molecule type: protein  
 A:Residues: 2-5, 'X', 7-11, 59-64, 125-132, 222-226, 'X', 228-230 <COL2>  
 C:Complex: heterotrimer of a 115K chain and two kinesin-related chains of 95K (PIR:S5869)  
 C:Superfamily: kinesin-related protein KIF3; kinesin motor domain homology  
 C:Keywords: ATP; heterotrimer; microtubule binding; nucleotide binding; P-loop  
 F:1-348/Domain: kinesin motor domain homology <KWOT>  
 F:97-104/Region: nucleotide-binding motif A (P-loop)  
 F:103/Binding site: ATP (lys) #status predicted

Query Match 37.0%; Score 623.5; DB 1; Length 699;  
 Best Local Similarity 46.2%; Pred. No. 9.4e-40;  
 Matches 156; Conservative 47; Mismatches 110; Indels 25; Gaps 10;

QY 6 VAVCVVRPLNSRRESLG-----ETAQVYWKTDNNVIYQVDGSKSFNDRVPHG 54  
 DB 11 VRVVRCPRLNSKETGQGFKSVVKMDMGTGVQV---TNPNA-PSGEPPKSTFDIVFAP 66  
 QY 55 NETTKNVYEIAPIIDSAIOQYNGTIFAYGQTASGKTYTMMG--SEDLH-GVIPRAIHD 111  
 DB 67 GAKQTDVYNTQARPIVDALIEGNGTIFAYGQTGKTPTMEGVRSPQELRGIIPNSFAH 126  
 QY 112 IFQIKKKFPDR-EFLLRVSYMEIYNETITDLCGTQKMKPLIREDVNRNVYADLTVEV 170  
 DB 127 IFGHIAKEQENRFLRVSYLEYINEEVKDLL-GKQQRHLEVKRDPDVGCVYKDLISAFV 185  
 QY 171 VYTSEALKWITTKGSKSRHYGETKMNORSRSHITFRMILESKRGPSCEGSKVYSHL 230  
 DB 186 VNNADDMDRIMTLGNKNRSVGA TNMNESSRSHTAIFITILERSDMG--LDKEQHVAVGKL 243  
 QY 231 NLVDLAGSRAAQTGAAGVRLKEGCNINRSFILGQVKKLSDGQVGGFNYRDSKLTRI 290  
 DB 244 HMDVLAGESEQTGTGQELKEATKINLSLTGLNVISSLDVGK-STHPIYRNSKLTRL 302  
 QY 291 LQNSLGNPKTRIICTTPV--SFDETLTALQFASAK 326  
 DB 303 LQDSLGNNAKTVMCANIGPAEYNYDETISTLRVANRAK 340

RESULT 14  
 T49235

kinesin-like protein - Arabidopsis thaliana  
 N:Alternate names: protein F7K15.60  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
 C:Accession: T49235  
 R:Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Rudd, S.; Lem  
 submitted to the Protein Sequence Database, April 2000  
 A:Reference number: Z25019  
 A:Accession: T49235  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-932 <OBE>  
 A:Cross-references: EMBL:AL353871; GSPDB:GN00061; ATSP:F7K15.60  
 A:Experimental source: cultivar Columbia; BAC clone F7K15  
 C:Genetics:  
 A:Gene: ATSP:F7K15.60  
 A:Map position: 3  
 A:Introns: 76/1; 106/1; 139/3; 177/3; 205/1; 227/3; 247/3; 289/2; 355/3; 401/3; 430/3;

Query Match 36.8%; Score 621; DB 2; Length 932;  
 Best Local Similarity 44.6%; Pred. No. 2.1e-39;  
 Matches 148; Conservative 54; Mismatches 110; Indels 20; Gaps 8;

QY 3 EGAVAVCVVRPLNSRRESLGSETAQVYWK--TDNNVIYQ----VDGSKSFNDRVPHGNE 56  
 DB 26 EEKILVTVMRPLNWRHAKYDL--IAWPCDDDTIVFKPNPDKAPTKYDFDKVFETC 83  
 QY 57 TTKNVYEIAPIIDSAIOQYNGTIFAYGQTASGKTYTMMGSEDLHGVIPRAIHDFOKI 116  
 DB 84 ATQEVYEGGGRDVALSALAGTNATIFAYGQTSSGKTFTMR-----GVTESWVKDIYBHI 137  
 QY 117 KKFDRPDRFLRVSYMEIYNETITDLCGTQKMKPLIREDVNRNVYADLTVEVYVVTSEM 176  
 DB 138 RKTQERSFLVKVSALEYINVTWDL--NRDTGLRLDDPEKGTIVENLVEEVVESKQH 195  
 QY 177 ALKWITTKGSKSRHYGETKMNORSRSHITFRMILESKRGPSCEGSKVYSHLNLVDLA 236  
 DB 196 LQHLISICEDQVQGETALNDKSRSHOIILTHSHSLR-ELAGCVQSF-MATLNLVDLA 253  
 QY 237 GSERAAQTGAAGVRLKEGCNINRSFILGQVKKLSDGQVGGFNYRNSKUTRILQNSLIG 296  
 DB 254 GSERAFQTNADGLRLKEGSHINRSLLTTLTVIRKLSGKRKHDPYRDSKLTRILQNSLIG 313

QY 297 GNPKTRIICTTPV--SFDETLTALQFASAK 326  
 DB 314 GNARTALICTISPALSHVEQTKKTLSPAMSAX 345

RESULT 15  
 T06065  
 hypothetical protein F19H22.150 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 31-Mar-2000  
 C:Accession: T06065  
 R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.  
 submitted to the Protein Sequence Database, March 1999  
 A:Reference number: Z15184  
 A:Accession: T06065  
 A:Molecule type: DNA  
 A:Residues: 1-121 <BEV>  
 A:Cross-references: EMBL:AL035679; GSPDB:GN00062; ATSP:F19H22.150  
 A:Experimental source: cultivar Columbia; BAC clone F19H22  
 C:Genetics:  
 A:Gene: ATSP:F19H22.150  
 A:Map position: 4  
 A:Introns: 139/2; 170/1; 200/1; 256/3; 302/3; 322/3; 349/1; 371/3; 390/3; 425/3;  
 C:Superfamily: kinesin heavy chain; kinesin motor domain homology  
 F:99-494/Domain: kinesin motor domain homology <KWOT>

Query Match 36.7%; Score 618.5; DB 2; Length 1121;  
 Best Local Similarity 39.4%; Pred. No. 4.3e-39;  
 Matches 158; Conservative 58; Mismatches 102; Indels 83; Gaps 12;





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 09:26:36 ; Search time 3.15048 Seconds  
(without alignments)

5421.082 Million cell updates/sec

Title: US-10-045-631B-88\_COPY\_2\_329

Perfect score: 1686

Sequence: 1 AEEGAVAVCVRVPLNSREE.....PVSFDETLALQFATAKYM 328

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1686	100.0	2663	1	CENE HUMAN
2	650	38.6	786	1	FL10_CHLRE
3	644	38.2	702	1	KF3A HUMAN
4	643	38.1	701	1	KF3A MOUSE
5	643	38.1	1232	1	KF4A HUMAN
6	642	38.1	742	1	KI21_STRPU
7	641.5	38.0	1231	1	KF4A_MOUSE
8	631	37.4	747	1	KF3B_HUMAN
9	630	37.4	747	1	KF3B_MOUSE
10	628.5	37.3	1226	1	KF4A_XENLA
11	624.5	37.0	1029	1	KI17_HUMAN
12	623.5	37.0	699	1	KI22_STRPU
13	613.5	36.4	1038	1	KI17_MOUSE
14	608	36.1	1225	1	KF4A_CHICK
15	605	35.9	672	1	OSN3_CAEEL
16	593.5	35.2	975	1	KINH_DROME
17	591.5	35.1	1031	1	KINH_STRPU
18	591.5	35.1	1584	1	U104_CAEEL
19	588	34.9	1749	1	KI3A_MOUSE
20	586.5	34.8	554	1	KLP3_SCHPO
21	585	34.7	793	1	KF3C_HUMAN
22	584	34.6	796	1	KF3C_RAT
23	584	34.6	1805	1	KI3A_HUMAN
24	583	34.6	963	1	KINH_MOUSE
25	582	34.5	796	1	KF3C_MOUSE
26	582	34.5	928	1	KINH_NEUCR
27	582	34.5	963	1	KINH_HUMAN
28	582	34.5	1066	1	KI61_DROME
29	580.5	34.4	1032	1	KINN_HUMAN
30	578.5	34.3	1027	1	KINN_MOUSE
31	578.5	34.3	1826	1	KI3B_HUMAN
32	578	34.3	957	1	KF5C_HUMAN
33	576.5	34.2	967	1	KINH_LOUPE

34 575.5 34.1 378 1 KLP2\_BOMMO  
35 574 34.0 956 1 KF5C\_MOUSE  
36 571.5 33.9 935 1 KINH\_SYNRA  
37 570 33.8 1690 1 KF1A\_HUMAN  
38 569 33.7 815 1 KINH\_CAEEL  
39 568 33.7 784 1 KLP8\_DROME  
40 568 33.7 1695 1 KF1A\_MOUSE  
41 561 33.3 1648 1 KF14\_HUMAN  
42 557 33.0 1816 1 KF1B\_MOUSE  
43 547 32.4 1103 1 KF1C\_HUMAN  
44 539.5 32.0 1056 1 KI25\_ARATH  
45 538 31.9 1067 1 BG52\_XENLA

P46874 bombyx mori  
P28738 mus musculus  
Q43093 synccephalas  
Q12756 homo sapien  
P34540 caenorhabdi  
P46867 drosophila  
P33173 mus musculus  
Q15058 homo sapien  
Q60575 mus musculus  
O43896 homo sapien  
P82266 arabidopsis  
Q91783 xenopus lae

## ALIGNMENTS

RESULT 1  
CENE HUMAN  
ID CENE HUMAN STANDARD; PRT; 2663 AA.  
AC Q02224;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Centromeric protein E (CENP-E protein).  
GN CENPE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93024922; PubMed=1406971;  
RA Yen T.J., Li G., Schaar B.T., Szilak I., Cleveland D.W.;  
RT "CENP-E is a putative kinetochore motor that accumulates just before  
RT mitosis.";  
RL Nature 359:536-539(1992).  
RN [2]  
RP CHARACTERIZATION.  
RX MEDLINE=95196755; PubMed=7889940;  
RA Thrower D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.;  
RT "Mitotic HeLa cells contain a CENP-E-associated minus end-directed  
RT microtubule motor.";  
RN EMBO J. 14:918-926(1995).  
RP CHARACTERIZATION.  
RX MEDLINE=98437347; PubMed=9763420;  
RA Chan G.K.T., Schaar B.T., Yen T.J.;  
RT "Characterization of the kinetochore binding domain of CENP-E reveals  
RT interactions with the kinetochore proteins CENP-F and hBUBR1.";  
RN J. Cell Biol. 143:49-63(1998).  
RP FARNESYLATION.  
RX MEDLINE=20459117; PubMed=10852915;  
RA Ashar H.R., James L., Gray K., Carr D., Black S., Armstrong L.,  
RA Bishop W.R., Kirschmeier P.;  
RT "Farnesyl transferase inhibitors block the farnesylation of CENP-E  
RT and CENP-F and alter the association of CENP-E with the  
RT microtubules.";  
RN J. Biol. Chem. 275:30451-30457(2000).  
CC -!- FUNCTION: MINUS-END DIRECTED MICROTUBULE MOTOR. PROBABLE  
CC KINETOCORE MOTOR. ACCUMULATES JUST BEFORE MITOSIS AT THE G2 PHASE  
CC OF THE CELL CYCLE. PROBABLY IMPORTANT FOR CHROMOSOME MOVEMENT  
CC AND/OR SPINDLE ELONGATION.  
CC -!- SUBUNIT: INTERACTS WITH CENP-F AND BUBR1 KINASE.  
CC -!- SUBCELLULAR LOCATION: ASSOCIATES WITH KINETOCHORES DURING  
CC CONGRESSION, RELOCATES TO THE SPINDLE MIDZONE AT ANAPHASE, AND IS  
CC QUANTITATIVELY DISCARDED AT THE END OF THE CELL DIVISION.  
CC -!- SIMILARITY: Belongs to the kinesin-like protein family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; Z15005; CAA78727.1; -.  
DR PIR; S28261; S28261.  
DR HSP; P17119; 3KAR.  
DR Genew; HGNC:1856; CENPE.  
DR GK; Q02224; -.  
DR MIM; 117143; -.  
DR GO; GO:0005699; C-kinetochore; TAS.  
DR GO; GO:0005634; C-nucleus; TAS.  
DR GO; GO:0008350; F-kinetochore motor activity; TAS.  
DR GO; GO:000067; P-DNA replication and chromosome cycle; TAS.  
DR GO; GO:0007079; P-mitotic chromosome movement; TAS.  
DR GO; GO:0007080; P-mitotic metaphase plate congression; TAS.  
DR InterPro; IPR001752; kinesin\_motor.  
DR Pfam; PF00225; kinesin; 1.  
DR PRINTS; PR00380; KINESINHEAVY.  
DR SMART; SM00129; KISC; 1.  
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.  
DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.  
KW Motor protein; Cell division; ATP-binding; Coiled coil; Mitosis;  
KW Cell cycle; Centromere; Lipoprotein; Prenylation.  
FT DOMAIN 1 335 KINESIN-MOTOR.  
FT DOMAIN 336 2471 COILED COIL (POTENTIAL).  
FT DOMAIN 2472 2663 GLOBULAR (POTENTIAL).  
FT NP\_BIND 86 93 ATP (BY SIMILARITY).  
FT LIPID 2660 2660 S-farnesyl cysteine.  
SQ SEQUENCE 2663 AA; 312087 MW; CECF1980C8C8CB8 CRC64;

Query Match 100.0%; Score 1686; DB 1; Length 2663;  
Best Local Similarity 100.0%; Pred. No. 2.5e-121;  
Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AEEGAVCVVRPLNSREESLGETAQYVWKTDNNVIYQVDGSKSEFDRVFGHNETKN 60  
Db 2 AEEGAVCVVRPLNSREESLGETAQYVWKTDNNVIYQVDGSKSEFDRVFGHNETKN 61  
Qy 61 VYEEIAPIIDSAIQGYNGTIFAYGQTASGKTYTMGSEDLGVIPRAIHDFQKIKKFP 120  
Db 62 VYEEIAPIIDSAIQGYNGTIFAYGQTASGKTYTMGSEDLGVIPRAIHDFQKIKKFP 121  
Qy 121 DREELLRVSYMEIYNETITDLCGTQKMKPLIREDVNRNRYVADLTVEVYVTSMAIKW 180  
Db 122 DREELLRVSYMEIYNETITDLCGTQKMKPLIREDVNRNRYVADLTVEVYVTSMAIKW 181  
Qy 181 ITGKESRHYGETKMNQSSRSHTIFRMLSRKGEPSNCEGSKVSHLNLVDLAGSER 240  
Db 182 ITGKESRHYGETKMNQSSRSHTIFRMLSRKGEPSNCEGSKVSHLNLVDLAGSER 241  
Qy 241 AAGTGAAGVRKKEGCNINRSLFILGVIKKLSGQGVGFINVRDLSKLTRILQNSLGGNPK 300  
Db 242 AAGTGAAGVRKKEGCNINRSLFILGVIKKLSGQGVGFINVRDLSKLTRILQNSLGGNPK 301  
Qy 301 TRIICTITPVSFDETLTALQFASTAKYM 328  
Db 302 TRIICTITPVSFDETLTALQFASTAKYM 329

RESULT 2  
FL10 CHLRE STANDARD; PRT; 786 AA.  
AC P46869;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DE 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Kinesin-like protein FLA10 (KHPI protein).  
GN FLA10.  
OS Chlamydomonas reinhardtii.  
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
OC Chlamydomonadaceae; Chlamydomonas.

OX NCBI\_TaxID=3055;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=137;  
RX MEDLINE=94299638; PubMed=8027176;  
RA Walther Z., Vashishtha M., Hall J.L.;  
RT "The Chlamydomonas FLA10 gene encodes a novel kinesin-homologous  
protein";  
RL J. Cell Biol. 126:175-188(1994).  
CC -!- FUNCTION: Probably involved in flagellar assembly and maintenance.  
CC May play a role in flagellar synthesis.  
CC -!- TISSUE SPECIFICITY: Flagellar axoneme.  
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN  
CC II SUBFAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; L33697; AAA21738.1; -.  
DR PIR; A53939; A53939.  
DR HSP; P17119; 3KAR.  
DR InterPro; IPR001752; kinesin\_motor.  
DR Pfam; PF00225; kinesin; 1.  
DR PRINTS; PR00380; KINESINHEAVY.  
DR SMART; SM00129; KISC; 1.  
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.  
DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.  
KW Motor protein; Microtubule; ATP-binding; Coiled coil.  
KW DOMAIN 1 358 KINESIN-MOTOR (BY SIMILARITY).  
FT DOMAIN 367 687 COILED COIL (POTENTIAL).  
FT DOMAIN 688 786 GLOBULAR (POTENTIAL).  
FT NP\_BIND 97 104 ATP (POTENTIAL).  
FT DOMAIN 388 391 POLY-GLY.  
FT DOMAIN 705 714 POLY-GLY.  
FT DOMAIN 756 759 POLY-ASP.  
SQ SEQUENCE 786 AA; 86671 MW; F90969203EB79F1B CRC64;

Query Match 38.6%; Score 650; DB 1; Length 786;  
Best Local Similarity 44.8%; Pred. No. 2.6e-42;  
Matches 154; Conservative 56; Mismatches 110; Indels 24; Gaps 8;

Qy 5 AVAVCVVRPLNSREESLGETAQYVWKTDNNVIYQVDGSKSEFDRVFGHNET 57  
Db 10 SVKVVVRCRPLNGKEKADGRSRIVDMVDAGQVVRNPKADASBPFAFTDQVYDNCQ 69  
Qy 58 TKNVYEEIAPIIDSAIQGYNGTIFAYGQTASGKTYTMGSEDLGVIPRAIHDFQ 114  
Db 70 QRDVFDITARPLIDSCIEGYNGTIFAYGQTGKSHTMGKDEPELRLGLIPNTPRYVFE 129  
Qy 115 KI-KKFPDRPRLRVSYMEIYNETITDLCGTQKMKPLIREDVNRNRYVADLTVEVYV 173  
Db 130 IIRDSGTKEELVRSSYLEIYNEEVRDL-LGKDHSGKMKELKESRPGVYVVDLSQFVCKN 188  
Qy 174 SEMALKWITKGEKSRHYGETKMNQSSRSHTIFRMLSRKGEPSNCEGSKVSHLNLVDLAGSER 224  
Db 189 YEEMNKVLLAGKDNKRVGATLMNQSSRSHTIFRMLSRKGEPSNCEGSKVSHLNLVDLAGSER 248  
Qy 225 VKVSHLNLVDLAGSERAAQTGAAGVRKKEGCNINRSLFILGVIKKLSGQGVGFINVRD 284  
Db 249 VRVGKLNLDLAGSERQKDTGATGDRKKEGINKLNLSLTALGNVISALVDGK-SGHI PYRD 307

Qy 285 SKLTRLQNSLGGNPKTRIICTITPV--SFDETLTALQFASTAK 326  
Db 308 SKLTRLQNSLGGNPKTRIICTITPV--SFDETLTALQFASTAK 351  
RESULT 3  
KF3A\_HUMAN



J. Cell Biol. 119:1287-1296(1992).

CC -!- FUNCTION: MICROTUBULE-BASED ANTEROGRADE TRANSLOCATOR FOR

CC MEMBRANOUS ORGANELLES. PLUS END-DIRECTED MICROTUBULE SLIDING

CC ACTIVITY IN VITRO.

CC -!- SUBUNIT: HETERODIMER OF KIF3A AND KIF3B.

CC -!- TISSUE SPECIFICITY: EXPRESSED ALMOST EXCLUSIVELY IN ADULT BRAIN

CC TISSUE (MAINLY IN THE CEREBELLAR GRANULAR LAYER) WITHIN A SINGLE

CC TYPE OF NEURONAL CELL.

CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN

CC II SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: D12645; BAA02166.1; -.

CC PIR: B44259; B44259.

CC HSP: P17119; 3KAR.

CC MGD; MGI:107689; Kif3a.

CC InterPro: IPR001752; kinesin\_motor.

CC Pfam: PF00225; kinesin; 1.

CC PRINTS; PR00380; KINESIN.

CC SMART; SM00129; KISC; 1.

CC PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.

CC PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.

CC Motor protein; Microtubule; ATP-binding; Coiled coil; Neurone.

CC FT DOMAIN 1 350 KINESIN MOTOR (BY SIMILARITY).

CC FT DOMAIN 351 598 COILED COIL (BY SIMILARITY).

CC FT NP BIND 100 107 ATP (BY SIMILARITY).

CC FT DOMAIN 442 445 POLY-GLU.

CC FT DOMAIN 509 512 POLY-ARG.

CC FT SEQUENCE 701 AA; 801.67 MW; 2405872DFD95A29 CRC64;

CC SQ

Query Match 38.1%; Score 643; DB 1; Length 701;

Best Local Similarity 46.2%; Pred. No. 7.6e-42;

Matches 154; Conservative 48; Mismatches 115; Indels 16; Gaps 8;

Qy 6 VAVCVVRPLNSRESL--GETAQYWKTDNNVYQVDGS----KSNFDRVEHGNETK 59

Db 15 KVWVRCPLNREKSMYQAVSDVDEMRGITVHKTDSSNEPKTFDVFQPEKQL 74

Qy 60 NYVEELAPITDSATQYNGTIFAYGQTASGKTYTMGSE---DHLGVIPRAIHDFOKI 116

Db 75 DVYNLTARPIIDSVLENGTIFAYGQTGKTFMEGVRAVPLGLGVIPNSFAHIFGHI 134

Qy 117 KKFP--DRFELRVSYMEIYNETITDLCGTQRMKPLIITRENVNRYVADLTFEVTYSE 175

Db 135 AKAEGRTRFLRVSYLEYNEVRDL--GKQDQRLVLRKPDVGVYIKDL SAYVNNAD 193

Qy 176 MALKWITGEKRRHGETKQNRSSRSRTIFPMILESREKGPSPCEGSKVYSHLNLDL 235

Db 194 DMDRIMTLGHKNRSGVATNMNEHSRSHAFITFIECKSGVDGNH--VRMGKHLVDL 251

Qy 236 AGSRAAOTGAAGVRLKEGCNINRSILFGLGVIKKLSGQGVGFINYRDSKLTRELNSL 295

Db 252 AGSEQAQKATGQRLKEATKLNLSLTGNVISALVDGK--STHVPYRNSKLTRELQPSL 310

Qy 296 GGNPKTRICTITPV--SPDETALALQFASTAK 326

Db 311 GGNSTKMMCANIGPADYNYDETISTILRYANRAK 343

RESULT 5

KF4A HUMAN

ID KF4A HUMAN STANDARD; PRT; 1232 AA.

AC Q95239; Q9NNY6; Q9NNY24; Q9UNW3;

DT 15-JUL-1999 (Rel. 38, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Chromosome-associated kinesin KIF4A (Chromokinesin).

GN KIF4A OR KIF4

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]\_TaxID=9606;

RN SEQUENCE FROM N.A.

RC TISSUE=Lymphocytes;

RA Villard L.;

RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RN SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.

RP MEDLINE=20435301; PubMed=10978527;

RX Oh S.J., Hain H., Torrey T.A., Shin H., Choi W., Lee Y.M.,

RA Morse H.C. III, Kim W.;

RT "Identification of the human homologue of mouse KIF4, a kinesin

RT superfamily motor protein.";

RT Biochim. Biophys. Acta 1493:219-224 (2000).

RN [3]

RN SEQUENCE FROM N.A.

RP TISSUE=Retinoblastoma;

RC Rentsch A., Neumann T., Rommerskirch W.;

RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RN [4]

RN SEQUENCE OF 128-1232 FROM N.A.

RP TISSUE=Retinoblastoma;

RX MEDLINE=97311419; PubMed=9168136;

RA Yan R.-T., Wang S.-Z.;

RT "Increased chromokinesin immunoreactivity in retinoblastoma cells.";

RT Gene 189:263-267 (1997).

CC -!- FUNCTION: REQUIRED FOR MITOTIC CHROMOSOMAL POSITIONING AND BIPOLAR

CC SPINDLE STABILIZATION (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Nuclear. Associated with mitotic

CC chromosomes (By similarity).

CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEMATOPOIETIC TISSUES,

CC FETAL LIVER, SPLEEN, THYMUS AND ADULT THYMUS AND BONE MARROW.

CC LOWER LEVELS ARE FOUND IN HEART, TESTIS, KIDNEY, COLON AND LUNG.

CC -!- SIMILARITY: Belongs to the kinesin-like protein family.

CC Chromokinesin subfamily.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: AF179308; AAD51855.1; -.

CC EMBL: AF071592; AAD05492.2; -.

CC EMBL: AJ271784; CAB75427.1; -.

CC EMBL: AF277375; AAF86334.1; -.

CC HSP: P17119; 3KAR.

CC Genew; HGNC:13339; KIF4A.

CC GO; GO:0005737; C:cytoplasm; TAS.

CC GO; GO:0005876; C:spindle microtubule; TAS.

CC GO; GO:0003777; F:microtubule motor activity; TAS.

CC GO; GO:0008089; P:anterograde axon cargo transport; TAS.

CC GO; GO:0006996; P:organellar organization and biogenesis; TAS.

CC InterPro: IPR001752; kinesin\_motor.

CC Pfam: PF00225; kinesin; 1.

CC PRINTS; PR00380; KINESIN.

CC SMART; SM00129; KISC; 1.

CC PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.

CC PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.

CC Motor protein; Microtubule; ATP-binding; DNA-binding;

CC Nuclear protein; Coiled coil.

CC FT DOMAIN 1 349 KINESIN-MOTOR.

CC FT DOMAIN 350 999 COILED COIL (BY SIMILARITY).

CC FT DOMAIN 1000 1232 GLOBULAR.

CC NP\_BIND 88 ATP (POTENTIAL).



FT CONFLICT 223 233 R -> G (IN REF. 2).  
 FT CONFLICT 231 231 S -> T (IN REF. 4).  
 FT CONFLICT 286 286 V -> A (IN REF. 2).  
 FT CONFLICT 422 422 L -> W (IN REF. 2).  
 FT CONFLICT 564 564 L -> H (IN REF. 4).  
 FT CONFLICT 564 564 L -> P (IN REF. 2).  
 FT CONFLICT 600 600 K -> E (IN REF. 3).  
 FT CONFLICT 668 668 R -> K (IN REF. 3 AND 4).  
 FT CONFLICT 928 928 Q -> P (IN REF. 1).  
 FT CONFLICT 958 958 Q -> R (IN REF. 3).  
 FT CONFLICT 960 960 L -> Q (IN REF. 1).  
 FT CONFLICT 996 997 LL -> S (IN REF. 4).  
 FT CONFLICT 1014 1014 QKHLPKDTLLSP -> RLPRIIPFYLO (IN REF. 4).  
 FT CONFLICT 1022 1022 P -> Q (IN REF. 2).  
 FT CONFLICT 1077 1077 K -> N (IN REF. 2).  
 FT CONFLICT 1138 1138 G -> S (IN REF. 2).  
 SQ SEQUENCE 1232 AA; 139908 MW; FF74052A17A8B8F7 CRC64;

Query Match 38.1%; Score 643; DB 1; Length 1232;  
 Best Local Similarity 44.4%; Pred. No. 1.6e-41;  
 Matches 147; Conservative 51; Mismatches 117; Indels 16; Gaps 5;

QY 6 VAVCVVRPLNSREESLG-ETAQVYKTDNNVIQVVDGSKSENFDFVFGHGNFTKNVYEE 64  
 Db 10 VRVALRCRPLVPKEISGQWCLSFVPGEPQVVGTD--KSFYDFVDFDPSTQEVRVNT 67  
 QY 65 IAPAIIDSAIQNGNFIAYGTQASGKTYTMG-----SEDLGVIPRAIHDFOKIK 117  
 Db 68 AVAPLKGKFGKYNATVLAAYGTGSGKTYSMGAYTAQENETVGVIPRVIQLLKEID 127  
 QY 118 KFPDFEFLRVSMYELNETITDLCGTQMKPLIREDVNRNYYVADLTREVVYTSMA 177  
 Db 128 KKSDFEFLRVSMYELNETITDLCGTQMKPLIREDVNRNYYVADLTREVVYTSMA 177  
 QY 178 LKWTGKESRHYGETKMQRSRSHITFRMILESEKGEPSNCEGSKVKVSHLNLVDLAG 237  
 Db 188 VSCLEGGNSRIVASTAMNSQSSRSHAITISLEQRKSD----KNSFRSKLHLVDLAG 243  
 QY 238 SERAQGTGAAGVRLKEGCNINSLFVLGVIKLSDGVGGFVINYRDSKLTILQNSLGG 297  
 Db 244 SERQKTKAEGDRLEKGININRGLICLVNIGALGDDKGGFVYRDSKLTILQNSLGG 303  
 QY 298 NPKTRIICITPV--SFDLTLALQFASAK 326  
 Db 304 NSHTLMIACVSPADNSLEETLTLRYADRAR 334

## RESULT 6

K121\_STRPU  
 ID K121\_STRPU STANDARD; PRT; 742 AA.  
 AC P46871;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DE 28-FEB-2003 (Rel. 41, Last annotation update)  
 DB Kinesin-II 95 kDa subunit (KRP-85/95 95 kDa subunit).  
 GN KRP95.  
 OS Strongylocentrotus purpuratus (Purple sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 OC Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;  
 OC Strongylocentrotus.  
 OX NCBI\_TaxID=7668;  
 RN [1]

## SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

TSUSS-EGG;  
 RC MEDLINE=94050179; PubMed=8232586;  
 EX Cole D.G., Chinn S.W., Wedaman K.P., Hall K., Vuong T.,  
 RA Scholey J.M.;  
 RT "Novel heterotrimeric kinesin-related protein purified from sea  
 RT urchin eggs";  
 RL Nature 366:268-270(1993).  
 CC -/- SUBUNIT: Heterotrimer of a 115 kDa subunit (KAP115) and two  
 CC kinesin-like subunits of 95 kDa (KRP95) and 85 kDa (KRP85).  
 CC -/- PTM: The N-terminus is blocked.

CC -/- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN  
 CC II SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; U00996; AAA87393.1; -;  
 DR HSSP; P17119; 3KAR.  
 DR InterPro; IPR001752; kinesin\_motor.  
 DR Pfam; PF00225; kinesin; 1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.  
 DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.  
 KW Motor protein; Microtubule; ATP-binding; Coiled coil.  
 FT DOMAIN 1 337 KINESIN-MOTOR (BY SIMILARITY).  
 FT DOMAIN 338 613 COILED COIL (BY SIMILARITY).  
 FT DOMAIN 614 742 GLOBULAR (BY SIMILARITY).  
 FT NP BIND 95 102 ATP (POTENTIAL).  
 SQ SEQUENCE 742 AA; 84202 MW; 47C40A367BAA77B5 CRC64;

Query Match 38.1%; Score 642; DB 1; Length 742;  
 Best Local Similarity 45.0%; Pred. No. 9.8e-42;  
 Matches 150; Conservative 50; Mismatches 117; Indels 16; Gaps 6;

QY 6 VAVCVVRPLNSREESLG-ETAQVYKTDNNVIQV-----DGSKSENFDFVFGHNEIT 58  
 Db 9 VKVVRCRPMNSKEISQGHKRVEMDNKRLGVETNPKGPPGPKNSKFTFDVYDWSKQ 68  
 QY 59 KNVYEIAPIIDSAIQNGNFIAYGTQASGKTYTMG---SEDLGVIPRAIHDFOK 115  
 Db 69 IDLYDETFSLVSVLQGGNGFIAYGTGTGKTFTMEGVRNPELGVIPNSFHFIFTH 128  
 QY 116 IKKFPDFEFLRVSMYELNETITDLCGTQMKPLIREDVNRNYYVADLTREVVYTS 175  
 Db 129 IARTQNGQFLVRASYLEIVQEIRDLAKDQK-KELDLKERPDVTGVYKDLSSFTVSKV 187  
 QY 176 LKWTGKESRHYGETKMQRSRSHITFRMILESEKGEPSNCEGSKVKVSHLNLVDL 235  
 Db 188 EIEHVTVGNNSRIVASTAMNSQSSRSHAITISLECELVDG--ENHIVGKLNLDL 245  
 QY 236 AGSERAQGTGAAGVRLKEGCNINSLFVLGVIKLSDGVGGFVINYRDSKLTILQNSL 295  
 Db 246 AGSERQAKTGATGDELKATKINLSLALGNVISALVDGK-SSHIPYRDSKLTILQNSL 304  
 QY 296 GGNPKTRIICITPV--SFDLTLALQFASAK 326  
 Db 305 GGNKTVVANNGPASYNFDEITITLRYANRAK 337

## RESULT 7

## KF4A\_MOUSE

ID KF4A\_MOUSE STANDARD; PRT; 1231 AA.  
 AC P33174;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Chromosome-associated kinesin KIF4A (Chromokinesin).  
 GN KIF4A OR KIF4 OR KNS4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE  
 RP SPECIFICITY.  
 RC STRAIN=ICR; TISSUE=Brain;  
 RX MEDLINE=95014709; PubMed=7929562;

RA Sekine Y., Okada Y., Noda Y., Kondo S., Aizawa H., Takemura R.,  
RA Hirokawa N.;  
RA "A novel microtubule-based motor protein (KIF4) for organelle  
RT transports, whose expression is regulated developmentally";  
RL J. Cell Biol. 127:187-201(1994).  
RN [2]  
RP SEQUENCE OF 91-240 FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=93077686; PubMed=1447303;  
RA Aizawa H., Sekine Y., Takemura R., Zhang Z., Nangaku M.,  
RA Hirokawa N.;  
RT "Kinesin family in murine central nervous system.";  
RL J. Cell Biol. 119:1287-1296(1992).  
CC -!- FUNCTION: REQUIRED FOR MITOTIC CHROMOSOMAL POSITIONING AND BIPOLAR  
CC SPINDLE STABILIZATION. Nuclear. Associated with mitotic  
CC CHROMOSOMES.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN PYRAMIDAL CELLS IN JUVENILE  
CC HIPPOCAMPUS, GRANULAR CELLS IN JUVENILE CEREBELLAR CORTEX AND IN  
CC ADULT SPLEEN.  
CC -!- SIMILARITY: Belongs to the kinesin-like protein family.  
CC Chromokinesin subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D12646; BAA02167.1; -  
CC PIR; A54803; A54803.  
CC HSSP; P17119; 3KAR.  
CC MGI; 108389; Kif4.  
CC InterPro; IPR001752; kinesin\_motor.  
CC Pfam; PF00225; kinesin; 1.  
CC PRINTS; PR00380; KINESINHEAVY.  
CC SMART; SM00129; KISC; 1.  
CC PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
CC PROSITE; PS00667; KINESIN\_MOTOR\_DOMAIN2; 1.  
CC Motor protein; Microtubule; ATP-binding; DNA-binding;  
CC Nuclear protein; Coiled coil.  
CC KINESIN-MOTOR.  
CC COILED COIL (BY SIMILARITY).  
CC GLOBULAR.  
CC NP\_BIND 88 95  
CC ATP (POTENTIAL).  
CC I -> S (IN REP. 2).  
CC CONFLICT 112 112  
CC SEQUENCE 1231 AA; 139551 MW; F34FZC2D21158FE4 CRC64;  
CC -----  
CC Query Match 38.0%; Score 641.5; DB 1; Length 1231;  
CC Best Local Similarity 44.3%; Pred. No. 2.1e-41;  
CC Matches 147; Conservative 53; Mismatches 115; Indels 17; Gaps 7;  
CC -----  
CC 6 VAVCVVRPLNREESLG-ETAQVYWKTNVYQVDSKSNFDFVFGHNETTKNVEE 64  
CC 10 VRVALRCPLVSKKEGCGTCLSFVPEQPQV--VGNDSFTYDFDPSTPEQEVNT 67  
CC 65 IAAPTDSALQNGTIFAYGQTASGTYTMGS---EDH---LGVIPRAIHDFQKIK 117  
CC 68 AVAPLKGVPKGNATVLAGYQSGKTYSMGAYTAQEHDASGIVIPRVIQLLFEIN 127  
CC 118 KFPDREFLRLVSYMEIYNITDLCCT-QKMKPLIREDVNRNRYVADITVEVYVSEM 176  
CC 128 KKSDFEFLKVSYLEIYNEIBLLCSREKATQINIREDPKGIKIVGLTEKVLVAD 187  
CC 177 ALKWIITKGEKSRHYGTYKMNQSRSHITFRMILESEKGEPCNCEGSVKVSHNLVDLA 236  
CC 188 TVSCLEQGNNSRTVASTANNQSQRSHAITFISIEQKK-----NDKNSFRSKLHVDLA 243  
CC 237 GSERAQGTGAAGVRLKGCNINRSLFLTGIVIKLSIQGVGFNRYRDSKILTRILNSLG 296  
CC 244 GSERQKTKAEGDLREGININRGLLCLGNVISAIGDDKGNFVPYRDSKILTRILQSLG 303

QY 297 GNPETRIITCTPV--SPDETITAIQFASTAK 326  
DB 304 GNSHTLMIACVSPADSNLEETINTLYRADAR 335  
  

RESULT 8  
KF3B HUMAN STANDARD; PRT; 747 AA.  
ID KF3B HUMAN  
AC O15066;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Kinesin-like protein KIF3B (Microtubule plus end-directed kinesin  
DE motor 3B) (HH0048).  
GN KIF3B OR KIAA0359.  
OS Homo sapiens (Human).  
OC Rukaryota; Metazoa;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=9734994; PubMed=9205841;  
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,  
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. VII.  
RT The complete sequences of 100 new cDNA clones from brain which can  
RT code for large proteins in vitro.";  
RL DNA Res. 4:141-150(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Clee C.M.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Cleghon S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Leheslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
RN [3]  
RP IDENTIFICATION IN A COMPLEX WITH SMC3 AND KIFAP3B.  
RX MEDLINE=98175913; PubMed=9506951;  
RA Shimizu K., Shirataki H., Honda T., Minami S., Takai Y.;  
RT "Complex formation of SNAP/KAP3, a KIF3A/B ATPase motor-associated  
RT protein, with a human chromosome-associated polypeptide.";  
RL J. Biol. Chem. 273:6591-6594(1998).  
CC -!- FUNCTION: Involved in tethering the chromosomes to the spindle  
CC pole and in chromosome movement. Microtubule-based anterograde  
CC translocator for membranous organelles. Plus end-directed  
CC microtubule sliding activity in vitro (By similarity).  
CC -!- SUBUNIT: Heterodimer of KIF3A and KIF3B (By similarity). Interacts  
CC with the SMC3 subunit of the cohesin complex.

```

CC CC      -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC CC      II SUBFAMILY.
CC CC      -----
CC CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC      the European Bioinformatics Institute. There are no restrictions on its
CC CC      use by non-profit institutions as long as its content is in no way
CC CC      modified and this statement is not removed. Usage by and for commercial
CC CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC      or send an email to license@isb-sib.ch).
CC CC      -----
CC CC      ENBL; AB002357; BAA20815.1; -
CC CC      ENBL; AL121897; CAC16425.1; -
CC CC      HSP; P17119; 3KAR.
CC CC      Genew; HGNC:6320; KIF3B.
CC CC      MIM; 603754; -.
CC CC      GO; GO:0005873; C:plus-end kinesin complex; TAS.
CC CC      GO; GO:0003777; F:microtubule motor activity; TAS.
CC CC      GO; GO:0008574; F:plus-end-directed kinesin ATPase activity; TAS.
CC CC      GO; GO:0008089; P:anterograde axon cargo transport; TAS.
CC CC      GO; GO:0007368; P:determination of left/right asymmetry; TAS.
CC CC      InterPro; IPR001752; kinesin_motor.
CC CC      Pfam; PF00225; kinesin; 1.
CC CC      PRINTS; PR00380; KINESINHEAVY.
CC CC      SMART; SM00129; KISC; 1.
CC CC      PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
CC CC      PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.
CC CC      Motor protein; Microtubule; ATP-binding; Coiled coil; Neurone.
CC CC      FT DOMAIN 1 345 KINESIN-MOTOR (BY SIMILARITY).
CC CC      FT DOMAIN 346 579 COILED COIL (BY SIMILARITY).
CC CC      FT NP BIND 580 747 GLOBULAR.
CC CC      FT DOMAIN 96 103 ATP (POTENTIAL).
CC CC      FT DOMAIN 386 393 POLY-GLY.
CC CC      FT DOMAIN 394 406 POLY-GLU.
CC CC      FT DOMAIN 723 730 POLY-SER.
CC CC      SQ SEQUENCE 747 AA; 85125 MW; 97FA4573AFA87023 CRC64;

Query Match
Best Local Similarity 37.4%; Score 631; DB 1; Length 747;
Matches 147; Conservative 54; Mismatches 115; Indels 20; Gaps 7;

QY 5 AVAVCVVRPLNSREESLG-----ETAQVYVKTDDNNVIYQVDGSKSFNDRVFHGN 55
Db 9 SVRVVVRCPWNGCKEAKAAYDKVDVDVKLGQSVKNPKGTASHEM--PKTFTFDAVDWN 66
QY 56 ETKNVYEEAAPIIDSAIQYNGTIFAYGQTASGKTYTM---MGSEDLHLGVIPRAIHI 112
Db 67 AKQFELYDETFRPLVDVSLQGFNGTIFAYGQTGKTYYTMGIRGDPKRGVIPSFDHI 126
QY 113 FQKIKKFPDRFLLRVSYMEIYNETITDLCGTQRMKPLIIRDVNRNVYADLTVEEVY 172
Db 127 FTHISRSQOQYLVRASYLEIYQEEIRDLSSKDQ--TKRLELKERPDGTGVYVKDLSFVTK 185
QY 173 TSEALKWITTKGKSRHYGETKMNORSSRSHITFRMILESREKGEPSNCGSVKSHLNL 232
Db 186 SVKEIEHVMVNGNQNSVGVATNMEHSSRSHAFVITIECEVSG--LDGENHIRVGKLN 243
QY 233 VDLAGSRAAGTGAAGVRLKEGCNINRSLIFILGQVKKLSDGVGVGFNYRDSKLTLQ 292
Db 244 VDLAGSERQAKTGAQGERLKEATKINLSLALGNVVISALVDGK--STHIPYRDSKLTLQ 302
QY 293 NSLGGNPKTRICHTITVSPF--DETILALQFASAK 326
Db 303 DSLGGNKTVMVNVGFSYNNVEETLTTLRYANRAK 338

RESULT 9
KF3B_MOUSE
ID_KF3B_MOUSE
AC_Q61711;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

```

```

DE DE      Kinesin-like protein KIF3B (Microtubule plus end-directed kinesin
DE DE      motor 3B).
DE GN      KIF3B.
DE OS      Mus musculus (Mouse).
DE OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DE OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
DE OX      NCBI_TaxID=10090;
DE RN      [1]
DE RP      SEQUENCE FROM N.A.
DE RC      STRAIN=ICR; TISSUE=Brain;
DE RX      MEDLINE=96032268; PubMed=7559760;
DE RA      Yamazaki H., Nakata T., Okada Y., Hirokawa N.;
DE RT      "KIF3A/B: a heterodimeric kinesin superfamily protein that works as a
DE RT      microtubule plus end-directed motor for membrane organelle
DE RT      transport.";
DE RL      J. Cell Biol. 130:1387-1399(1995).
DE CC      -!- FUNCTION: Involved in tethering the chromosomes to the spindle
DE CC      pole and in chromosome movement. Microtubule-based anterograde
DE CC      translocator for membranous organelles. Plus end-directed
DE CC      microtubule sliding activity in vitro (By similarity).
DE CC      -!- SUBUNIT: Interacts with the SMC3 subunit of the cohesin
DE CC      complex (By similarity). Heterodimer of KIF3A and KIF3B.
DE CC      -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
DE CC      II SUBFAMILY.
DE CC      -----
DE CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
DE CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
DE CC      the European Bioinformatics Institute. There are no restrictions on its
DE CC      use by non-profit institutions as long as its content is in no way
DE CC      modified and this statement is not removed. Usage by and for commercial
DE CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
DE CC      or send an email to license@isb-sib.ch).
DE CC      -----
DE DR      EMBL; D26077; BAA05070.1; -.
DE DR      PIR; A57107; A57107.
DE DR      HSP; P17119; 3KAR.
DE DR      MGD; MGI:107688; Kif3b.
DE DR      InterPro; IPR001752; kinesin_motor.
DE DR      Pfam; PF00225; kinesin; 1.
DE DR      PRINTS; PR00380; KINESINHEAVY.
DE DR      SMART; SM00129; KISC; 1.
DE DR      PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DE DR      PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.
DE KW      Motor protein; Microtubule; ATP-binding; Coiled coil; Neurone.
DE FT      DOMAIN 1 345 KINESIN-MOTOR (BY SIMILARITY).
DE FT      DOMAIN 346 579 COILED COIL (BY SIMILARITY).
DE FT      NP BIND 580 747 GLOBULAR.
DE FT      DOMAIN 96 103 ATP (POTENTIAL).
DE FT      DOMAIN 386 393 POLY-GLY.
DE FT      DOMAIN 394 405 POLY-SER.
DE FT      DOMAIN 723 730 POLY-SER.
DE SQ      SEQUENCE 747 AA; 85288 MW; FA369A4190ECB47 CRC64;

Query Match
Best Local Similarity 37.4%; Score 630; DB 1; Length 747;
Matches 147; Conservative 54; Mismatches 115; Indels 20; Gaps 7;

QY 5 AVAVCVVRPLNSREESLG-----ETAQVYVKTDDNNVIYQVDGSKSFNDRVFHGN 55
Db 9 SVRVVVRCPWNGCKEAKAAYDKVDVDVKLGQSVKNPKGTASHEM--PKTFTFDAVDWN 66
QY 56 ETKNVYEEAAPIIDSAIQYNGTIFAYGQTASGKTYTM---MGSEDLHLGVIPRAIHI 112
Db 67 AKQFELYDETFRPLVDVSLQGFNGTIFAYGQTGKTYYTMGIRGDPKRGVIPSFDHI 126
QY 113 FQKIKKFPDRFLLRVSYMEIYNETITDLCGTQRMKPLIIRDVNRNVYADLTVEEVY 172
Db 127 FTHISRSQOQYLVRASYLEIYQEEIRDLSSKDQ--TKRLELKERPDGTGVYVKDLSFVTK 185
QY 173 TSEALKWITTKGKSRHYGETKMNORSSRSHITFRMILESREKGEPSNCGSVKSHLNL 232
Db 186 SVKEIEHVMVNGNQNSVGVATNMEHSSRSHAFVITIECEVSG--LDGENHIRVGKLN 243

```

QY 233 VDLGASERAAOTGAAGVRLKEGCNINRSIFILGQVKKLSDGQVGFNRYRDSKLTIRLQ 292  
 DB 244 VDLASERCAKGAQGERLKEATKINLSLALGNVISALVDGK-SNTHIPYRDSKLTIRLQ 302

QY 293 NSLGNPKTRIICTIPVSF--DETALQFASAK 326  
 DB 303 DSLGGNAKTVMANVGPASYNVEETLTTRYANRAK 338

RESULT 10  
 ID KF4A XENLA STANDARD; PRT; 1226 AA.  
 AC Q91784; Q9PS10;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Chromosome-associated kinesin KLP1 (Chromokinesin).  
 GN KLP1.

OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN NCBI\_TaxID=8355;  
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.

RC TISSUE=Oocyte;  
 RX MEDLINE=95236444; PubMed=7720067;  
 RA Vernos I., Raats J., Hirano T., Heasman J., Karsenti E., Wylie C.;  
 RT "Xklp1, a chromosomal Xenopus kinesin-like protein essential for spindle organization and chromosome positioning.";  
 RL Cell 81:117-127(1995).  
 RN [2]

RP SEQUENCE OF 9-338 FROM N.A.  
 RX MEDLINE=93246065; PubMed=8482413;  
 RA Vernos I., Heasman J., Wylie C.;  
 RT "Multiple kinesin-like transcripts in Xenopus oocytes.";  
 RL Dev. Biol. 157:232-239(1993).  
 CC -!- FUNCTION: REQUIRED FOR MITOTIC CHROMOSOMAL POSITIONING AND BIPOLAR  
 CC -!- SPINDLE STABILIZATION.

CC -!- SUBCELLULAR LOCATION: Nuclear. Associated with mitotic chromosomes.  
 CC -!- TISSUE SPECIFICITY: Expressed in oocytes, eggs, testes and brain.  
 CC -!- SIMILARITY: Belongs to the kinesin-like protein family.  
 CC Chromokinesin subfamily.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X82012; CAA57539.1; -.  
 DR PIR; I51617; I51617.  
 DR HSSP; P17119; 3KAR.  
 DR InterPro; IPR001752; kinesin\_motor.  
 DR PRINTS; PF00225; kinesin; 1.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.  
 DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.  
 KW Motor protein; Microtubule; ATP-binding; DNA-binding;  
 KW Nuclear protein; Coiled coil.  
 FT DOMAIN 1 350 KINESIN-MOTOR.  
 FT DOMAIN 351 1006 COILED COIL (BY SIMILARITY).  
 FT DOMAIN 1007 1226 GLOBULAR.  
 FT NP BIND 87 94 ATP (POTENTIAL).  
 FT CONFLICT 163 163 I -> L (IN REF. 2).  
 SQ SEQUENCE 1226 AA; 138923 MW; 7F0275FCF3316697 CRC64;

Query Match 37.3%; Score 628.5; DB 1; Length 1226;  
 Best Local Similarity 44.1%; Pred. No. 2.1e+04;  
 Matches 149; Conservative 53; Mismatches 117; Indels 19; Gaps 9;  
 QY 2 EGG-AVAVCVRVPLNSREESLG-ETAQVYWKTDNNVYQVDGSKSFNDRFVHGNETTK 59  
 DB 4 DEGIPIRVVALRCPLVPKENNEGCKMCLTFVPEQQVI--VGTESFTDYVDFPSAEQE 61  
 QY 60 NYVEETAAPLIDSAIQVNGTIFAYGQTASGKTYTMMGSEDH-----LGVIPRAIHDI 112  
 DB 62 EVYNSAVAPLIKGLFKGYNATVLAQGTGSGKTYSGGAYTHNQENPTVGVIPRTVIAL 121  
 QY 113 FQIKKFPDPREFLLRVSYMEINNETIDLL-CGTQMKELIREDVNRNVYVADLTVEEV 171  
 DB 122 FREIHPQPEWENFLKVSYLEIYNEEILDLAARDKNTTISREDKEGKIKGLTERDV 181  
 QY 172 YTSEMALKWITKEKSRHYGETKNQORSSRSHTIFRMILESRKBPSPNCEGSKVSHLN 231  
 DB 182 KTAALDTLSCLQGNSSRTVASTAMNSQSSRSHAFITISIQRKEGKNN--SFR-SKLH 237  
 QY 232 LVDLGASERAAOTGAAGVRLKEGCNINRSIFILGQVKKLSD-GQVGGFNYRDSKLTIRI 290  
 DB 238 LVDLGASERQKTKYAGEDRLKEGINSRGLLCGNVISALGDESCKGFPYRDSKLTIRL 297  
 QY 291 LQNSLGNPKTRIICTIPV--SFDETALQFASAK 326  
 DB 298 LQDSLGNSTLMTACVSPADSNMEETLTRYADRAR 335

RESULT 11

ID KF17 HUMAN STANDARD; PRT; 1029 AA.  
 AC Q9P2E2; Q95077; Q8N411;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Kinesin-like protein KIF17 (KIF3-related motor protein).  
 GN KIF17 OR KIF3X OR KIAA1405.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kimberley A., White S.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 78-1029 FROM N.A. (ISOFORM 1).  
 RC TISSUE=Brain;  
 RA Ohara O., Nagase T., Kikuno R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 239-1029 FROM N.A. (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=20181126; PubMed=10718198;  
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro.";  
 RL DNA Res. 7:65-73(2000).  
 RN [4]  
 RP SEQUENCE OF 618-1029 FROM N.A. (ISOFORM 2).  
 RC TISSUE=Pancreas;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K., Altschul S.F., Zeeberg B., Buetow K.H., Wang J., Hsieh F., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stappleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,



FT DOMAIN 1 340 KINESIN-MOTOR (BY SIMILARITY).  
 FT DOMAIN 341 619 COILED COIL (BY SIMILARITY).  
 FT DOMAIN 620 699 GLOBULAR (BY SIMILARITY).  
 FT NP\_BIND 97 104 ATP (POTENTIAL).  
 SQ SEQUENCE 699 AA; 78697 MW; 7B3866111CB08190 CRC64;  
 Query Match 37.0%; Score 623.5; DB 1; Length 699;  
 Best Local Similarity 46.2%; Pred. No. 2.4e-40;  
 Matches 156; Conservative 47; Mismatches 110; Indels 25; Gaps 10;  
 QY 6 VAVCVVRPLNREBSLG-----ETAQVYWKTDNNVYQVDSKSNFDRVFG 54  
 DB 11 VRVVRCPNLSKETCGPKSVVMDMRGTQV-----TPNA-PSEPPEKSFDFVFP 66  
 QY 55 NETTKNVEELAPIIDSIAIQYNGTIFAYGQTASGKTYTMMG--SEDHL-GVIPRAIHD 111  
 DB 67 GAKQTVYVQNTARPIVDALIEGNGTIFAYGQTGKTFMEGVRSQPELRGIIPNSFAH 126  
 QY 112 IFQIKKFPDR-EFLRVSMYIYNETITDLCGTQKMKPLIREDVNRNVYADLVEEV 170  
 DB 127 IFGHIAKEQENRFLVRVSYLEIYNEEVKDLL-GKQOHRLEVKRPPDVGVIYVYKDJSAFV 185  
 QY 171 VYTSMAKWKWTGKSRHYGETKMKORSRSHITFERMILESRKGEPSNCEGSKVYVSHL 230  
 DB 186 VNNADMDRIMTLGNKNRSVGATNMNESSRSHAFTITILERSDMG--LDKEQHVVRGKL 243  
 QY 231 NLVDLAGSRAAQTGAAGVRLKKEGNCINRSLFILGQVVKKLSGQGVGFYNYRDSKLTRI 290  
 DB 244 HMVDLAGSEKQTKGATGQRLKEATKINLSLSTLGNVISLVDGK-STHPIYRNSKLTRL 302  
 QY 291 LQNSLGNPKTRICITTPV--SFDETLTALQFPASTAK 326  
 DB 303 LQDSLGNNAKTVMCANIGPAEYNYDETISTLRYANRAK 340  
 RESULT 13  
 KF17 MOUSE ID KF17 MOUSE STANDARD; PRT; 1038 AA.  
 AC Q99P8;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Kinesin-like protein Kif17 (MmKif17).  
 GN KIF17.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20307907; PubMed=10846156;  
 RA Setou M., Nakagawa T., Seog D.-H., Hirokawa N.;  
 RT "Kinesin superfamily motor protein Kif17 and mLin-10 in NMDA  
 receptor-containing vesicle transport.";  
 RL Science 288:1796-1802 (2000).  
 CC -!- FUNCTION: Transports vesicles containing N-methyl-D-aspartate  
 (NMDA) receptor 2B along microtubules.  
 CC -!- SUBUNIT: Interacts with LIN-10 PDZ domain.  
 CC -!- TISSUE SPECIFICITY: Neuronal-specific.  
 CC -!- SIMILARITY: Belongs to the kinesin-like protein family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; AB008867; BAB21099.1; -.  
 DR HSSP; PF17119; 3KAP.  
 DR MGD; MGI:1098229; Kif17.

DR GO; GO:0005871; C:kinesin complex; IDA.  
 DR GO; GO:0003777; F:microtubule motor activity; IDA.  
 DR GO; GO:0005515; F:protein binding; IPI.  
 DR GO; GO:0007017; P:microtubule-based process; IDA.  
 DR GO; GO:0016192; P:vesicle-mediated transport; IDA.  
 DR InterPro; IPR001752; P:vesicle-mediated transport; IDA.  
 DR Pfam; PF00225; kinesin; 1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.  
 DR PROSITE; PS00677; KINESIN MOTOR DOMAIN2; 1.  
 KW Motor protein; Microtubule; ATP-binding; Coiled coil; Transport;  
 KW Protein transport.  
 FT DOMAIN 1 265 KINESIN-MOTOR.  
 FT DOMAIN 346 470 COILED COIL (POTENTIAL).  
 FT DOMAIN 748 855 COILED COIL (POTENTIAL).  
 FT NP\_BIND 91 98 ATP (POTENTIAL).  
 SQ SEQUENCE 1038 AA; 116372 MW; 2BED852A3AFBD46 CRC64;  
 Query Match 36.4%; Score 613.5; DB 1; Length 1038;  
 Best Local Similarity 42.3%; Pred. No. 2.4e-39;  
 Matches 142; Conservative 57; Mismatches 116; Indels 21; Gaps 8;  
 QY 5 AVAVCVVRPLNREBSLG-ETAQVYWKTDNNVYQVDS-----KSNFDRVPHGNETT 58  
 DB 5 SVKVVVRCRPMNKRRELSQSVVTVDSARGCQCFIQNPGAADPPKQFTFDGAYYIEHFT 64  
 QY 59 KNYVEETAPIIDSAIQYNGTIFAYGQTASGKTYTMMGSED---HLGVIPRAIHDFQK 115  
 DB 65 EQIYEIAYPLVEGVTEGNGTIFAYGQTGSGKSFYMQGLPDPQCQGIIPRAPEHFVES 124  
 QY 116 IKKFPDREFLLRVSMYIYNETITDLCGTQKMKPLIREDVNRNVYADLVEEVYTS 175  
 DB 125 VQCAENTKFLVRASYLEIYNEVDHLL-GADTKQRLKEHPEKGVYVKGSLSMHTVHVA 183  
 QY 176 MALKWITGKSRHYGETKMKORSRSHITFERMILE---SREKGEPSNCEGSKVYVSHL 232  
 DB 184 QCVRMETGWKNRAVGYTLMNKDSRSHSFTINIEIYAYVDERG-----KHLRAGKLN 238  
 QY 233 VDLAGSRAAQTGAAGVRLKKEGNCINRSLFILGQVVKKLSGQGVGFYNYRDSKLTILQ 292  
 DB 239 VDLAGSRSQTKGATGRLKEATKINLSLGNVISALVDGRC-KHPIYRDSKLTILQ 297  
 QY 293 NSLGNPKTRICITTPV--SFDETLTALQFPASTAK 326  
 DB 298 DSLGNTKTLMVACLSPADNNYDETISTLRYANRAK 333  
 RESULT 14  
 KF4A CHICK ID KF4A CHICK STANDARD; PRT; 1225 AA.  
 AC Q90640; Q90608;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Chromosome-associated kinesin KIF4A (Chromokinesin).  
 GN KIF4A.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE  
 RP SPECIFICITY.  
 RC STRAIN=White leghorn; TISSUE=Embryonic retina;  
 RX MEDLINE=95181533; PubMed=7876303;  
 RA Wang S.Z., Adler R.;  
 RT "Chromokinesin: a DNA-binding, kinesin-like nuclear protein.";  
 RL J. Cell Biol. 128:761-768 (1995).  
 RN [2]  
 RN SEQUENCE OF 728-1088 FROM N.A.  
 RC STRAIN=White leghorn; TISSUE=Embryonic retina;



```

RX MEDLINE=94151328; PubMed=8108415;
RA Wang S.Z., Adler R.;
RT "A developmentally regulated basic-leucine zipper-like gene and its
RL expression in embryonic retina and lens.";
Proc. Natl. Acad. Sci. U.S.A. 91:1351-1355 (1994).
CC -!- FUNCTION: REQUIRED FOR MITOTIC CHROMOSOMAL POSITIONING AND BIPOLAR
CC SPINDLE STABILIZATION.
CC -!- SUBCELLULAR LOCATION: Nuclear. Associated with mitotic
CC chromosomes.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN PROLIFERATING CELLS;
CC NEUROEPITHELIUM OF EMBRYOS.
CC -!- SIMILARITY: Belongs to the kinesin-like protein family.
CC Chromokinesin subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U18309; AAC59666.1; -.
DR EMBL; U04821; AAL18960.1; -.
DR PIR; A56514; A56514.
DR HSSP; P17119; 3KAR.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
KW Motor protein; Microtubule; ATP-binding; DNA-binding;
KW Nuclear protein; Coiled coil.
FT DOMAIN 1 351 KINESIN-MOTOR.
FT DOMAIN 352 1003 COILED COIL (BY SIMILARITY).
FT DOMAIN 1004 1225 GLOBULAR.
FT NP BIND 98 95
FT CONFLICT 1087 1088 KG -> RI (IN REF. 2).
FT SEQUENCE 1225 AA; 138923 MW; FA01ED83425F5875 CRC64;
Query Match 36.1%; Score 608; DB 1; Length 1225;
Best Local Similarity 42.4%; Pred. No. 8e-39;
Matches 143; Conservative 58; Mismatches 118; Indels 18; Gaps 8;
QY 2 EGAVAVCVRVPLNSREESLG-ETAQVYVKKTDNNVIVQVDSKGFNDRVPHGNETKN 60
Db 6 EKGIPVRVRCRPLVPKETSQCQCLSFVPGEPQV--VGSDKAFTYDYVDPDSVEQEE 63
QY 61 VVEETAAPIIDSAIQYNGTIFAYQGTASGKTYTMMG----SEHD---LGVIPRAIHDF 113
Db 64 VNTAVAPLIRGIFKGYNATVLAQGTGSGKTYSMGGTYTASQEHDPMSGVIPRVIKLLF 123
QY 114 QKIKKFPDREELLRVSYMEIYNETITDLCCT-QKMKPLIREDVNRNVYADLITEEVY 172
Db 124 KEKQRQDWEFVLKVSLEYLEINEDILLDCSSRSQSISREDPKGKIGVGLTERNVA 183
QY 173 TSEMAKWKITKGEKSRHYGETKMNQSRSSHTIFRMILESEKGEPSKVKVSHLNT 232
Db 184 SARDTVSCLEQNNCRVTASTAMNSQSSRSHAITICIDQKK-----NDKSSSFHKLHL 239
QY 233 VDLAGSERAATGAGVRLKGCNINRSLFLTGQVKKL-SDGVGGFINTYRDKLTRL 291
Db 240 VDLAGSERQKTKRAGEKRLKGININRGLLCLGNVISAALGKGNKGGFVPRYDLSKLTLL 299
QY 292 QNSLGGNPKTKIITCTTPV--SEDETALOPASTAK 326
Db 300 QDSLGGNSHTLMIACVSPADSNLETLNTRYADRA 336
RESULT 15
OSM3_CABEEL STANDARD; PRT; 672 AA.
ID -OSM3_CABEEL

```

```

AC P46873;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Kinesin-like protein osm-3.
GN OSM-3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodirinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=95230679; PubMed=7714894;
RA Tabish M., Siddiqui Z.K., Nishikawa K., Siddiqui S.S.;
RT "Exclusive expression of C. elegans osm-3 kinesin gene in
RT chemosensory neurons open to the external environment.";
J. Mol. Biol. 247:377-389 (1995).
RN [2]
RP PRELIMINARY SEQUENCE OF 1-397 FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=93379214; PubMed=7690265;
RA Shafir M.A., Fukushige T., Yasuda H., Miwa J., Siddiqui S.S.;
RT "C. elegans osm-3 gene mediating osmotic avoidance behaviour encodes
RT a kinesin-like protein.";
NeuroReport 4:891-894 (1993).
RL NeuroReport 4:891-894 (1993).
CC -!- TISSUE SPECIFICITY: AMPHID AND IL2 NEURONS.
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC II SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D38632; BAA07612.1; -.
DR EMBL; D14968; BAA20996.1; -.
DR PIR; S54351; S54351.
DR HSSP; P17119; 3KAR.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
KW Motor protein; Microtubule; ATP-binding; Coiled coil.
FT DOMAIN 57 300 KINESIN-MOTOR.
FT NP BIND 461 497 COILED COIL (POTENTIAL).
FT NP BIND 130 137 ATP (POTENTIAL).
FT SEQUENCE 672 AA; 75456 MW; 4E2160F7042AFCD7 CRC64;
Query Match 35.9%; Score 605; DB 1; Length 672;
Best Local Similarity 46.9%; Pred. No. 6e-39;
Matches 137; Conservative 47; Mismatches 94; Indels 14; Gaps 6;
QY 41 DG-SKSFNFRVPHGNETKNVVEEIAAPIDSALQYNGTIFAYQGTASGKTYTMMGSE 99
Db 85 DGAARDTFDGAQVPMIRPGQIVNDIVFLENVIEGNGTVFAYGTSGKTFMQGIE 144
QY 100 D---HLGVIPRAIHDFQKIKKFPDREELLRVSYMEIYNETITDLCCTQKMKPLIRED 156
Db 145 TIPAQRGVIPRAFDHIFTATATTENKVLKVSLEYLEINEVEVDLLGADNKKQ-LEIKQ 203
QY 157 VNRNVYADLITEEVYTSSEMAKWKITKGEKSRHYGETKMNQSRSSHTIFRMILESEK 216
Db 204 PDRGVYVAGLSMEVCHDVPACKELMTRGFNNRVHGATLMNKDSSRSHSITFVVEGITE- 262
QY 217 EPSNCEGSKVKVSHLNVDLAGSERAATGAGVRLKGCNINRSLFLTGQVKKLSDGV 276
Db 263 -----TGSIRMGKLNLDLAGSERQSKTGATGRLKEATKINLSLGAIGNVIALVDGK- 316

```

us-10-045-631b-88\_copy\_2\_329.rsp

Mon Aug 9 11:02:49 2004

QY 277 GGFINYRDSKLTIRILONSLGGNPKTRIICTITPVS--FDETLTALQFASTAK 326  
DB 317 SKHIPYRDSKLTIRILQDSLGGNPKTIRIMTACVSPSSDNYDETLTSTLRANRAK 368

Search completed: July 29, 2004, 09:36:13  
Job time : 4.15048 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 09:27:31 ; Search time 13.9321 Seconds  
(without alignments)  
7428.155 Million cell updates/sec

Title: US-10-045-631b-88\_COPY\_2\_329

Perfect score: 1686

Sequence: 1 AEGAVAVCVRVPLNSREE.....PVSFDETLALQFASTAKYM 328

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1485	88.1	549	11 Q7TPX4	Q7tpx4 mus musculus
2	1258.5	74.6	2954	13 Q42263	Q42263 xenopus lae
3	773.5	45.9	807	10 Q94HV9	Q94hv9 arabidopsis
4	773.5	45.9	823	10 Q957P3	Q957p3 arabidopsis
5	744.5	44.2	459	10 Q9SS30	Q9ss30 arabidopsis
6	738	43.8	888	10 Q9LQ62	Q9lq62 arabidopsis
7	719	42.6	1885	5 Q869B8	Q869b8 dictyosteli
8	707	41.9	160	11 Q35059	Q35059 mus musculus
9	688	40.8	2013	5 Q9VKI0	Q9vki0 drosophila
10	688	40.8	2244	5 Q9NCG0	Q9ncg0 drosophila
11	683	40.5	1055	10 Q8RWM4	Q8rwm4 arabidopsis
12	683	40.5	1055	10 Q8W5R5	Q8w5r5 arabidopsis
13	673	39.9	1058	10 Q9SUJ0	Q9suj0 arabidopsis
14	667.5	39.6	1033	10 Q9LHL9	Q9lhl9 arabidopsis
15	663.5	39.4	1459	3 P87198	P87198 ustilago ma
16	656	38.9	890	10 Q8W5R6	Q8w5r6 arabidopsis

17	654.5	38.8	954	10 Q9AWM8	Q9awm8 oryza sativ
18	652	38.7	629	5 Q8MTF8	Q8mpt8 caenorhabdi
19	649	38.5	959	10 Q8S950	Q8s950 nicotiana t
20	646.5	38.3	956	10 Q9C7B9	Q9c7b9 arabidopsis
21	644	38.2	408	4 Q8IWH8	Q8iwh8 homo sapien
22	643	38.1	408	11 Q8OUK1	Q8ouk1 mus musculu
23	643	38.1	443	11 Q8CGJ1	Q8cgj1 mus musculu
24	643	38.1	671	4 Q86XX7	Q86xx7 homo sapien
25	643	38.1	701	11 Q7TSZ7	Q7tsz7 mus musculu
26	643	38.1	1127	4 Q86TN3	Q86tn3 homo sapien
27	641.5	38.0	819	11 Q7TQG6	Q7tqg6 mus musculu
28	641.5	38.0	997	10 Q8FG03	Q8f903 arabidopsis
29	641.5	38.0	1231	11 Q80YP3	Q80yp3 mus musculu
30	639	37.9	699	13 Q98T11	Q98t11 xenopus lae
31	635.5	37.7	735	5 Q9UUD5	Q9u0d5 tetrahymena
32	631.5	37.5	974	10 Q8S905	Q8s905 arabidopsis
33	630.5	37.4	909	10 Q9X103	Q9x103 arabidopsis
34	630	37.4	747	11 Q8BNH4	Q8bnh4 mus musculu
35	630	37.4	757	11 Q8OU27	Q8ou27 mus musculu
36	629.5	37.3	836	5 Q7YUC7	Q7yuc7 tetrahymena
37	629.5	37.3	1193	10 Q7X7H8	Q7x7h8 oryza sativ
38	627.5	37.2	677	5 Q9VRK9	Q9vrk9 drosophila
39	621	36.8	932	10 Q9LXL3	Q9lxl3 arabidopsis
40	621	36.8	937	10 Q8LGU3	Q8lgu3 arabidopsis
41	621	36.8	937	10 Q8L5J2	Q8l5j2 arabidopsis
42	621	36.8	938	10 Q8LNZ2	Q8lnz2 arabidopsis
43	618.5	36.7	1121	10 Q9SVI8	Q9svi8 arabidopsis
44	617	36.6	671	5 Q8NPT7	Q8npt7 caenorhabdi
45	617	36.6	672	5 Q9GV93	Q9gv93 caenorhabdi

#### ALIGNMENTS

#### RESULT 1

Q7TPX4	PRELIMINARY;	PRT;	549 AA.
ID	Q7TPX4		
AC	Q7TPX4;		
DT	01-OCT-2003 (TREMBLrel. 25, Created)		
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)		
DE	01-OCT-2003 (TREMBLrel. 25, Last annotation update)		
DE	Hypothetical protein (Fragment).		
OS	Mus musculus (Mouse)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Egg;		
RX	MEDLINE=22388257; PubMed=12477932;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.R.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Krzywinski M.I., Skalski U., Smailus D.E., Schnerch A., Schein J.E.,		
RA	Jones S.J., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Egg;		



```

Db      55 TNASVYELLTDIIHAAVEGNGTAFAYGQTSSGKTTMTGSETDPGIIIRSRVDVFERI 114
QY      117 KFPDREFLRVSMYMEIYNETITDLCCTQKMKPLIREDVNRNVYVADLTETEEVVYTTSEM 176
Db      115 HMTSDREFLRVSMYMEIYNEEINDLL--AVENQLQIHEHLERGCVFVAGLKEEIVSDAEQ 172
QY      177 ALKWIITGKESRHYGETKMKQSRSHRTIFRMILESEKGEPCNCEGSKVSHNLNLDLA 236
Db      173 ILKLIDSGEVNRHFGETNMNVHSRSHRTIFRMVIESR--GKDNSSSDAIRVSVLNLVDLA 230
QY      237 GSERAQTGAAGVRLKEGCNINRSLFLGVIVIKLSDG--QVGGFINVRDSDKLTILQNSL 295
Db      231 GSRIAKTAGGVRLQEGKYINKSLMLGNVINKLSDTKLRAHIPYRDSKLTILQPAL 290
QY      296 GGNPKTRIITCTTPVS--FDETLTALQFASTAK 326
Db      291 GGNAKTCTIITAPEEHIEESKGTILQFASRAK 323

RESULT 4
Q9S7P3
ID Q9S7P3 PRELIMINARY; PRT; 823 AA.
AC Q9S7P3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Kinesin-like protein.
GN ZCF125.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Ooe H., Kato A., Komeda Y.;
RT "Arabidopsis Thaliana genomic sequence for a kinesin-like protein.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kato A., Suzuki M., Kuwahara A., Ooe H., Higano-Inaba K., Komeda Y.;
RT "Isolation and analysis of cDNA within a 300 kb Arabidopsis thaliana
RT genomic region located around the 100 map unit of chromosome 1.";
RL Gene 239:309-316(1999).
DR EMBL; AB028470; BAA88114.1; -
DR EMBL; AB028468; BAA88112.1; -
DR PIR; T52425; T52425.
DR HSRF; P33176; 1BG2.
DR GO; GO:0005871; C:kinesin complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0007017; P:microtubule-based process; IEA.
DR InterPro; IPR001752; P:microtubule-based process; IEA.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
SQ SEQUENCE 823 AA; 93148 MW; 6AFB1C622E4632C9 CRC64;

Query Match 45.9%; Score 773.5; DB 10; Length 823;
Best Local Similarity 52.0%; Pred. No. 1.5e-52;
Matches 173; Conservative 46; Mismatches 89; Indels 25; Gaps 7;

QY      6 VAVCVVRPLNRSRESIGETAQVYWKVDNNVIYQVDSKSFN-----FDRVFHNE 56
Db      4 ICVAVRVRP-----PAPENGASLWKVEDN---RISLHKSIDTPTTASHAFDHFDESS 54
QY      57 TTKNVYEIAAPITDSALOGYNGTIFAYGQTASGKTYTMGSEDLGVIPRAIHIFOKI 116

```

```

Db      55 TNASVYELLTDIIHAAVEGNGTAFAYGQTSSGKTTMTGSETDPGIIIRSRVDVFERI 114
QY      117 KFPDREFLRVSMYMEIYNETITDLCCTQKMKPLIREDVNRNVYVADLTETEEVVYTTSEM 176
Db      115 HMTSDREFLRVSMYMEIYNEEINDLL--AVENQLQIHEHLERGCVFVAGLKEEIVSDAEQ 172
QY      177 ALKWIITGKESRHYGETKMKQSRSHRTIFRMILESEKGEPCNCEGSKVSHNLNLDLA 236
Db      173 ILKLIDSGEVNRHFGETNMNVHSRSHRTIFRMVIESR--GKDNSSSDAIRVSVLNLVDLA 230
QY      237 GSERAQTGAAGVRLKEGCNINRSLFLGVIVIKLSDG--QVGGFINVRDSDKLTILQNSL 295
Db      231 GSRIAKTAGGVRLQEGKYINKSLMLGNVINKLSDTKLRAHIPYRDSKLTILQPAL 290
QY      296 GGNPKTRIITCTTPVS--FDETLTALQFASTAK 326
Db      291 GGNAKTCTIITAPEEHIEESKGTILQFASRAK 323

RESULT 5
Q9S830
ID Q9S830 PRELIMINARY; PRT; 459 AA.
AC Q9S830;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative kinesin-like centromere protein.
GN F14P13.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA Roming C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC F14P13 genomic sequence.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC009400; AAF02823.1; -
DR HSSP; P33176; 1BG2.
DR GO; GO:0005871; C:kinesin complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0007017; P:microtubule-based process; IEA.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
SQ SEQUENCE 459 AA; 51872 MW; 7C8487E9B7038E6A CRC64;

Query Match 44.2%; Score 744.5; DB 10; Length 459;
Best Local Similarity 56.5%; Pred. No. 1.3e-50;
Matches 161; Conservative 42; Mismatches 71; Indels 11; Gaps 5;

QY      49 DRVPHGNETTKNVYEETAAPIIDSAIOGYNGTIFAYGQTASGKTYTMGSEDLGVIPRA 108
Db      45 DRIFREDCKTVQVYEATKTEIVSAVRGNGTVPAYGQTNSGKTHTRGSGPIEFGLA 104
QY      109 IHDIFQIKFPDRREFLLRVSMEIYNETITDLCCTQKMKPLIREDVNRNVYVADLTETEE 168
Db      105 VHDLFDTIYQDASREFLLRVSMEIYNEEDINDLL--APEHRKLOIHENLEKGFVAGLRE 162
QY      169 EVVYTTSEALKWITKGEKSHYGETKMKQSRSHRTIFRMILESEKGEPCNCEGSKVSHNLNLDLA 224
Db      163 EIVASPOQVLEMMFEGESHRHIGETNMNLYSSRSHRTIFRMVIESR--GKDNSSSDAIRVSVLNLVDLA 221
QY      225 VKYSHNLNLDLAGSRAAQTGAAGVRLKEGCNINRSLFLGVIVIKLSDG--QVGGFINVRDSDKLTILQPAL 282

```

58 EIVSLFLFSFGVVVFLLFADHVFDESNASVYELLTKDIIHAAVEGNGTAFAYGQTS 117  
89 SKKTYMGSSEHLGVIPIRAHDIQKIKKFPDRFLLRVSVMEIYNETITDLCCTQKX 148  
118 SKKTYMGSSEHLGVIPIRAHDIQKIKKFPDRFLLRVSVMEIYNETITDLCCTQKX 175  
149 KPLIIREDVNRNVVADITVEEVVYVSEMAKWIITKEKSRHYGETONORSSRSHTIFRM 208  
176 QRLQHEHLERGVFVAGLKEEIVSDAEQILKIDSSEVNRHFGETNMVHSSRSHTIFRM 235  
209 -----ILRESKEGPEPNCSEGVSVKSHLNLVDLAGSRAAQTGAAGVRLKEGCN 256  
236 VRFSSYERDILLVIESR--GKNSSSDAIRVSVNLVLDLAGSRIAKTGAGGVRLQEGKY 293  
257 INRSLFLLGVVKKLSDG-QVGGFINYRDSKLTIRLQNSLGGNPKTRICTTTPVS--FD 313  
294 INKSLMILGNVINKUSDSKTLRAHIPYRDSKLTIRLQNSLGGNPKTRICTTTPVS--FD 353  
314 ETLTALQFASAK 326  
354 ESKGTQLQFASRAK 366

RESULT 7  
Q869B8 PRELIMINARY; PRT; 1885 AA.  
ID Q869B8  
AC Q869B8  
DT 01-JUN-2003 (TEMBLrel. 24, Created)  
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
DE Kinesin-related protein K4.  
GN K4.  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
OX NCBI\_TaxID=44689;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9839834; PubMed=9693369;  
RA de Hostos E.L., McCaffrey G., Sugang R., Pierce D.W., Vale R.D.;  
RT "A developmentally regulated kinesin-related motor proteins from  
RT Dictyostelium discoideum";  
RL Mol. Biol. Cell 9:2093-2106 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX Suyama E., Sutoh K.;  
RA "Kinesin-related proteins from Dictyostelium";  
RT Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AB102780; C:membrane; IEA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005875; C:microtubule associated complex; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0003774; F:motor activity; IEA.  
DR InterPro; IPR001752; kinesin motor.  
DR InterPro; IPR003345; M\_repeat.  
DR Pfam; PF00225; kinesin; 1.  
DR Pfam; PF02370; M; 6.  
DR PRINTS; PR00380; KINESINHEAVY.  
DR SMART; SM00129; KISC; 1.  
DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
DR PROSITE; PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.  
SQ SEQUENCE 1885 AA; 218216 MW; 15686AED3B007BC7 CRC64;

Query Match 42.6%; Score 719; DB 5; Length 1885;  
Best Local Similarity 48.8%; Pred. No. 1e-47;  
Matches 161; Conservative 56; Mismatches 101; Indels 12; Gaps 6;

QY 2 EGAVAVCVVRPLNSREESLGTAQVYKTDNNVI-YQVDGSKSFNDRVPHGNETKN 60  
DB ELNKIKVAIRVPLNSRE--LGIDQKIPWSISKDTISLQSNENINFTYDVGIDNTID 76  
QY VVEETAAPTIDSAIQNGTIFAYGQTASGKTYTMMGSEDLGLVTPRAIHDFQKIK-KF 119  
DB VYNALAKSLVNSLNGTIFAYGQTSKGKTFMRGTSIFGLIKLSIKDFKSIEDSI 136

222 VRVSVNLVLDLAGSRAAKTGAEGVRLKEGSHNKSMLTGLTVIKKLSGVTQGHVPI 281  
283 RDSKLTIRLQNSLGGNPKTRICTT--PVSEFDETLTALQFASTA 325  
282 RDSKLTIRLQNSLGGNPKTRICTT--PVSEFDETLTALQFASTA 326

Q9LQ62 PRELIMINARY; PRT; 888 AA.  
ID Q9LQ62  
AC Q9LQ62  
DT 01-OCT-2000 (TEMBLrel. 15, Created)  
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
DE T30E16.9  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX Chao Q., Brooks S., Buehler E., Dunn P., Khan S., Kim C., Shinn P.,  
RA Altati H., Araujo R., Conn L., Conway A.B., Gonzalez A., Hansen N.F.,  
RA Huizar L., Kremenetskaia I., Lenz C., Li J., Liu S., Luross S.,  
RA Rowley D., Schwartz J., Toriumi M., Vysotskaia V., Yu G., Davis R.W.,  
RA Federspiel N.A., Theologis A., Ecker J.R.;  
RT "Genomic sequence for Arabidopsis thaliana BAC T30E16 from chromosome  
RT I.";  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX Ecker J.R.;  
RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX Ecker J.R.;  
RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
RA Khan S., Kim C., Altati H., Bai B., Chin C., Chioi J., Choi E.,  
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,  
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsy N.,  
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
RA Theologis A., Ecker J.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AC009317; AAF9747.1; -.  
DR PIR; D96619; D96619.  
DR HSP; P33176; 1BG2.  
DR GO; GO:0005871; C:kinesin complex; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0003774; F:motor activity; IEA.  
DR GO; GO:0007017; F:microtubule-based process; IEA.  
DR InterPro; IPR001752; kinesin motor.  
DR Pfam; PF00225; kinesin; 1.  
DR PRINTS; PR00380; KINESINHEAVY.  
DR SMART; SM00129; KISC; 1.  
DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
DR PROSITE; PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.  
SQ SEQUENCE 888 AA; 100694 MW; 0D640FBACEE801B85 CRC64;

Query Match 43.8%; Score 738; DB 10; Length 888;  
Best Local Similarity 46.4%; Pred. No. 1.1e-49;  
Matches 173; Conservative 46; Mismatches 92; Indels 62; Gaps 8;

QY 6 VAVCVVRPLNSREESLGTAQVYKTDNNVI-----Y 38  
DB ICVAVVRP-----PAPENGASLWKVDNRNLSLHSLDTPITTHASHVSGISISTDLI 57  
QY 39 QVDGSKSFN-----DRVFHGNFTKNVYEEIAPIIDSAIQNGTIFAYGQTA 88

```
QY 120 PDRELLRVSYMEIYNEITDLCCT-QMKPLIIRVDNRNVVYADLTEEYVYTSMAL 178
Db 137 LEQDYLKVSYLEYNEIEIKOLLNPTISNKKKLIHEDIYKGVVANKKEIVISPDQIF 196
QY 179 KWTIKGEKSRHYGETKMNORSRSHITFRMILESEKGEPSNCEGSKVSHNLVDLAGS 238
Db 197 ALMNFGEERRHIGSTMWDSRSRSHITFRMIOISTCKQ-----NGTIQMSTLTVDLA 251
QY 239 BRAAQGTGAAGVRLKEGNCNINRSLFGLQVVKISDQGVGGFINVRDLSKLTILQNSL 298
Db 252 ERVSTGAEGVRLKEGTHINKSLMVSLSKSEKTKQHVPRDLSKLTILQPSLGGN 311
QY 299 PKTRILITIPVSP-DETTLALOPASTAK 326
Db 312 SKTAILCTITPATTHQBSISTLQFAKRAK 341

RESULT 8
O35059 PRELIMINARY; PRT; 160 AA.
AC O35059;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE Motor domain of KIF10 (Fragment).
GN CENPE OR KIF10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR;
RX MEDLINE=97420736; PubMed=92751178;
RA Nakagawa T., Tanaka Y., Matsuoaka E., Kondo S., Okada Y., Noda Y.,
RA Kanai Y., Hirokawa N.;
RT "Identification and classification of 16 new kinesin superfamily (KIF)
RT proteins in mouse genome."
RL Proc. Natl. Acad. Sci. U.S.A. 94:9654-9659 (1997).
DR EMBL; AB001426; BAA2386.1; -.
DR HSSP; P17119; 3KAR.
DR MGD; MGI:1098230; Cenepe.
DR GO; GO:0005871; C:kinesin complex; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0007017; P:microtubule-based process; IEA.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
FT NON_TER 1
FT NON_TER 160
SQ SEQUENCE 160 AA; 18406 MW; 9E6E4F6E2642C241 CRC64;

Query Match
Best Local Similarity 41.9%; Score 707; DB 11; Length 160;
Matches 135; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 81 IFAYGQTASGTYTMMGSEDLHGVIPRAIHDI FQIKKFPREFLLRVSYMEIYNEITD 140
Db 1 IFAYGQTASGTYTMMGSEDLHGVIPRAIHDI FQIKKFPREFLLRVSYMEIYNEITD 60
QY 141 LLCTQKMKPLIIRVDNRNVVYADLTEEYVYTSMALKWTIKGEKSRHYGETKMNORS 200
Db 61 LLCNAQKMKPLIIRVDNRNVVYADLTEEYVYTSMALKWTIKGEKSRHYGETKMNORS 120
QY 201 RSHTIFRMILSRKGPSPNCEGSKVSHNLVDLAGSE 239
Db 121 RSHTIFRMILSRKGPSPNCEGSKVSHNLVDLAGSE 159
```

```
RESULT 9
O3VK10 PRELIMINARY; PRT; 2013 AA.
AC O3VK10;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 16, Last sequence update)
DE Cmet protein.
DE Cmet OR CG6392.
GN Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.Y.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houstoun K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mohary C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
DR EMBL; AE003631; AAF53088.2; -.
DR HSSP; P33176; 1BG2.
DR FlyBase; FBgn0040232; cmet.
DR GO; GO:0005699; C:kinetochore; IEA.
DR GO; GO:0007080; P:mitotic metaphase plate congression; IMP.
DR InterPro; IPR001752; kinesin_motor.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR SMART; SM00150; SPEC; 3.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
SQ SEQUENCE 2013 AA; 231107 MW; 8A1A465581E4F5F CRC64;

Query Match
Best Local Similarity 40.8%; Score 688; DB 5; Length 2013;
Matches 135; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
```

Marches	162;	Conservative	48;	Mismatches	99;	Indels	16;	Gaps	8;
y	5	AVAVCVRPPLNSRESLSG	TAQVWYKT--DNNVYQV	DG-SKSFN	PDRVFHGN	ETTKNY	62		
b	8	SIQVCIKVRPCPG	LTSL-----WQVKERS	SIHLADSH	ARPYVDY	VDFECAS	NOEYF 60		
y	63	BEIAAPI	DSATQGVNGTTFAY	GOTASGTYT	TMGSEDL	GVPIRAIH	DFQIKKFPDR 122		
b	61	DRMAKIHVAC	QMGFGNGTTFAY	GOTSSGTYT	MMGDEQ	PNVULAAK	EIFQOISSETER 120		
y	123	EFLLRVS	MEIYNETTID	LCGTQK	WKPLIRED	VNRNVYAD	LTEVVYTSW-ALKWI 181		
b	121	DFLLRGY	IEIYNEKTYD	LL--NKKQ	DLKHESG	GINVNV--	CECIITSEVDLLRL 176		
y	182	TKGEKSRH	GETKMNORSRSH	TFPML	ESRKGE	PSNCEG	SVKVSMLNVL	LAGSERA 241	
b	177	CLGNKERT	VTGETNNKRS	SHAFKII	IESR-KSDHSD	--DDAVI	QSVLNLVL	LAGSERA 234	
y	242	AOTGAAG	VRILKEGCNTN	RSFLIL	GOVTKL	SDGOVG	GFNYRDSKL	TRILQNSL	GCNPKT 301
b	235	DQTGARG	ARKLKEGGHINK	SLLFL	SNVIKSL	SENADNR	FTNYRDSKL	TRILQASL	IGNAFT 294
y	302	RICTTP	VSFDETL	ALQAF	STAK	326			
b	295	SIICTKPS	IMEESQ	TLSPAT	RAK	319			

RESULT 10	
99NCGO	
DD	PRELIMINARY; PRT; 2244 AA.
Q9NCGO	
AC	
OT	01-OCT-2000 (TrEMBLrel. 15, Created)
OT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
OT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
OT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Kinesin-like kinetochore motor protein CENP-meta.
DE	CMT2 OR CG6392.
GN	Drosophila melanogaster (Fruit fly).
GN	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
OC	
OC	
OX	NCBI_TaxID=7227;
OX	[1]
RRN	SEQUENCE FROM N.A.
RRP	MEDLINE=20351410; PubMed=10893249;
RY	Yucel J.K., Marszalek J.D., McIntosh J.R., Goldstein L.S.B.,
RA	Cleveland D.W., Philip A.V.
RA	"CENP-meta, an Essential Kinetochore Kinesin Required for the
RA	Maintenance of Metaphase Chromosome Alignment in Drosophila.;"
RT	J. Cell Biol. 150:1-12(2000).
RL	EMBL; AF220353; AAF32355.1; -.
DR	HSP; P33176; 1BG2.
DR	FlyBase; FBgn0040232; cmet.
DR	GO; GO:0003699; C:kinetochore; IDA.
DR	GO; GO:0007080; P:mitotic metaphase plate congression; IMP.
DR	InterPro; IPR001752; kinesin_motor.
DR	Pfam; PF00225; kinesin; 1.
DR	PRINTS; PR00380; KINESINHEAVY.
DR	SMART; SM00129; K1SC; 1.
DR	PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR	PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
DR	SEQUENCE 2244 AA; 257992 MW; FA6AA3B2541ADE0 CRC64;
SO	

[illegible]

123 EFLLRYSYMEIYNETITDLLCGTQKMKPLIREDVNRRNVVADLTBEVVYITSEM-ALKW1 181  
 121 DFLLRYSYIIEYNEKIYDLL--NKKNQDLKHESGGNGIVNVN--CEECIITSEVDLRL 176  
 182 TKGEKSRHYGETKONORSSSHTIFRMILEBSRKEGPSNCEGSKVYSHLNLVDLAGSERA 241  
 177 CLGNKERTVGETNNRSRSHAFKIIIESR-KSDHSD-DDAVIDOSVLMVLVDLAGSERA 234  
 242 AQTCAAGVRLKEGGINRSLFILGOVIKKLSDGVGVGFINRYRDSKLTIRLONSLGNPKT 301  
 235 DOTGARGARLKEGGINKSLFLSNVYKLSLENADNFTNRYRDSKLTIRLOASLGNGAFT 294  
 302 RIICITITPVSDETLTALQFAS2AK 326  
 295 SIICITKPSIMEESQSTLSFATRAK 319  
 RESULT 11  
 QGRW4  
 DDYIMINADY. PRT: 1055 AA.

RESULT 11

Q8R9W4 PRELIMINARY; PRT; 1055 AA.

Q8R9W4; 21, Created)

01-JUN-2002 (TRENBLrel. 21, Last sequence update)

01-JUN-2002 (TRENBLrel. 21, Last sequence update)

01-OCT-2003 (TRENBLrel. 25, Last annotation update)

Putative kinesin.

AT4G39050.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eumetazoa II; Brassicales; Brassicaceae; Arabidopsids.

NCBI\_TaxID=3702;

[1]

SEQUENCE FROM N.A.

RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,

RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,

RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,

RA Carinci P., Chen H., Cheuk R., Hayaashiraki Y., Iehida J., Jones T.,

RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,

RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,

RA Sakurai T., Satou M., Seki M., Shimizu P., Southwick A., Shinozaki K.,

RA Davis R.W., Ecker J.R., Theologis A.,

RA "Arabidopsis Full Length cDNA Clones";

RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

RL EMBL; AY091060; AAM13881.1; -.

DR GO; GO:0005871; C:kinesin complex; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0003774; F:motor activity; IEA.

DR GO; GO:0007017; P:microtubule-based process; IEA.

DR InterPro; IPR001752; kinesin\_motor.

DR InterPro; IPR001841; Znf\_ring.

DR Pfam; PF00225; kinesin; 1.

DR PRINTS; PR00380; KINESINHEAVY.

DR SMART; SM00129; KISC; 1.

DR SMART; SM00184; RING; 1.

DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.

DR PROSITE; PSS0067; KINESIN MOTOR\_DOMAIN2; 1.

DR PROSITE; PS50089; ZF\_RING2; 1.

DR PROSITE; 1055 AA; 116462 NW; BBP3709ACEB40215B CRC64;

SO SEQUENCE

Query Match	40.5%;	Score 683;	DB 10;	Length 1055;
Best Local Similarity	47.6%;	Pred. No. 3.3e-45;		
				Gaps 6;

QY	1	AEEGAVACVVRPLNSREISLGTACVYKNTDNVYQVGDGSKSNFDFVPHGNEILAN	80
DB	94	SERDSISVTVFRPLSDREYQRQDEVAMYPDGLTVRHEYNPLTAYAFDKVFGPQATIID	153
QY	61	VYBEIAPIIDSAQVNGTITFAYGATSGKTYTVMGSEDLHGVIPRAIHDIFQIKKFP	120
DB	154	YDVNAARPVKARMEGVNGTVFAGVTSKGTHTMHGDSQSPGIPLAKDVFSTIOTPT	213
QY	121	DREFLLRVSYMEINETHITDLLCGTKMKPLIREDVNRYVYADLTSEVVYVTSEMAKLV	180





	DR	InterPro; IPR001841; Znf_ring.	
	DR	Pfam; PF00225; kinesin; 1.	
	DR	PRINTS; PR00380; KINCSINHEAVY.	
	DR	SMART; SM00129; KINCSIN MOTOR DOMAIN1; 1.	
	DR	SMART; SM00129; KISC; 1.	
	DR	SMART; SM00184; RING; 1.	
	DR	PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.	
	DR	PROSITE; PS00677; KINESIN MOTOR_DOMAIN2; 1.	
	DR	PROSITE; PS50089; ZF_RING_2; 1..	
	DR	SEQUENCE 1058 AA; 117089 MW; F97660F723A9A506 CRC64;	
	SQ		
	Query Match	39.9%; Score 673; DB 10; Length 1058;	
	Best Local Similarity	46.3%; Pred.No. 2e-44;	
	Matches 152; Conservative 57; Mismatches 109; Indels 10; Gaps 6;		
	Dr		
	Qy	1 AEGGAVCVRVPLNSREESLGETAQVYKTDNNVIYOVGDGSKSFNFRVFHGNETTKN 60	
	Dd	100 SERDSISVTFFRPFMSREQORDEIVTPDDAKVRNEYNPITAYAFKVFGPOSTTPE 159	
	Qy	61 VYEIAAPIIDSAIQGYNGTFIAYGTASGKYTMWGSDHLGVIPRAIHDFQIKKKFP 120	
	Dd	160 YVDVAAPVKVAAMEGVNTGFAYGVTSSGKTTHMGDQQDPGLIILAIDKFVSIOETT 219	
	Qy	121 DREFLLRSVMIEYNETHDLLCGTQMKNLIIREDVNRRNVYADLTERRVYTSEMALKW 180	
	Dd	220 GREFLLRVSYLEIYNEINDLPOTG--QNLRIRER-SQTYVEGEKEEVLSPGHALSF 276	
	Qy	181 ITKCEKRSHGETKMQRSSSRHTIFRMILESEKGPSCSGSVKSHLNLDLAGSR 240	
	Dd	277 IAAGEHHVGSNFNLMSSRESHTIFILMIESSAHGQDYD--GVIFPSQLNLIDLAGSE- 332	
	Qy	241 AAOATGAAGVLREGCNINRSFILGOVVKSLDGOGVGFNYRDSDKLTRLONSLGNPK 300	
	Dd	333 SSKTETTGLRKEGAYINKSLTLTGTVIGKLTEGKT-THPVPFRDSKLTRELQSLSGHGH 391	
	Qy	301 TRIICTITPV--SFDETALQFASTAK 326	
	Dd	392 VSLLCTVTPTASSSTEETHNLKFSARAK 419	
	RESULT 14		
	Q9LHL9	PRELIMINARY; PRT; 1033 AA.	
	ID		
	AC	Q9LHL9	
	DT	01-OCT-2000 (TREMBLrel. 15, Created)	
	DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
	DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)	
	DE	Kinesin (Centromere protein) like heavy chain-like protein.	
	OS	Arabidopsis thaliana (Mouse-ear cress).	
	OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
	OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;	
	OC	eurosid II; Brassicales; Brassicaceae; Arabidopsidis.	
	OX	NCBI_TaxID=3702;	
	RN	[1]	
	RP	SEQUENCE FROM N.A.	
	RC	STRAIN=Columbia;	
	RA	Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;	
	RL	Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.	
	RN	[2]	
	RP	SEQUENCE FROM N.A.	
	RC	STRAIN=Columbia;	
	RX	MEDLINE=20363099; PubMed=10907853;	
	RA	Nakamura Y.;	
	RT	"Structural analysis of Arabidopsis thaliana chromosome 3. II.	
	RT	Sequence features of the regions of 4,251,695 bp covered by ninety P1,	
	RT	TAC and BAC clones."	
	RL	DNA Res. 7:217-221(2000).	
	DL	EMBL; AP002040; BAB03114.1; --	
	DR	HSSP; P33176; 1BG2.	
	DR	GO; GO:0005871; C:kinesin complex; IEA.	
	DR	GO; GO:0005524; F:ATP binding; IEA.	
	DR	GO; GO:0003774; F:motor activity; IEA.	
	DR	GO; GO:0007017; F:microtubule-based process; IEA.	
	DR	InterPro; IPR001752; kinesin_motor.	
	Query Match	39.4%; Score 663.5; DB 3; Length 1459;	
	Best Local Similarity	47.3%; Pred.No. 5.4e-44;	
	Matches 158; Conservative 52; Mismatches 101; Indels 23; Gaps 10;		
	Dr		
	Qy	6 VAVCVRVRPLNSREESLGETAQVYKTDNNVIYQVGGSK--SFNDRVFHGNETTKNYE 63	
	Dd	70 VTTVTRPLSPREIRQE-EVAVYADGETVNRHNPTIAYAYDRVFGPTTTTRNVYD 127	
	Qy	64 ETAAPIIDSAIQGYNGTFIAYGTASGKYTMWGSDHLGVIPRAIHDFQIKKKFPDRE 123	
	Dd	128 IAAHVVMGAMGEGINGTIAYGVTSSGKTTHMHGDOGRSPCIIPLVAKDAFSIIQTENRE 187	
	Qy	124 FLHRVSMYELNYETIIDLLCGTQMKMP-----LIIRDVNRRNVYADLTERRVYTSEMALK 179	
	Dd	188 FLRIISWEIYNEVNDLL-----NPAGHNLIRED-KQGTVEGIKEEVLSPAHALS 240	
	Qy	180 WITKGE---KSRYHGETKMQRSSSRHTIFRMILESEKGPSCSGSVKSHLNLDLA 236	
	Dd	241 LIAAGEVMTQRHVGSFTNFLSSRSHTIFTLETIESPLGDKSGE-AVHLSQNLVDLA 299	
	Qy	237 GSBEAAQTGAAGVLREGCNINRSFILGOVVKSLDGOGVGFINTRDSDLTRLONSLG 296	
	Dd	300 GSE--SSKVTSGVVRRKEGYSINKSLTLTGTVISKLTDDR-AHSVYPYRDSKLTRELQSSL 357	
	Qy	297 GNPKTRIICITPV--SFDETALQFASTAKYM 328	
	Dd	358 GHDRVSLICTVTPTASSSTEETHNLKFAHRAKH 391	
	RESULT 15		
	P87198	PRELIMINARY; PRT; 1459 AA.	
	AC	P87198	
	DT	01-JUL-1997 (TREMBLrel. 04, Created)	
	DT	01-JUL-1997 (TREMBLrel. 04, Last sequence update)	
	DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)	
	DE	Kinesin motor protein.	
	GN	KINI.	
	OS	Ustilago maydis (Smut fungus).	
	OC	Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;	
	OC	Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.	
	OX	NCBI_TaxID=5270;	
	RN	[1]	
	RP	SEQUENCE FROM N.A.	
	RX	MEDLINE=97361828; PubMed=9218789;	
	RA	Lehmmer C., Steinberg G., Snetselaar K.M., Schliwa M., Kahmann R.,	
	RA	Bolker M.;	
	RT	"Identification of a motor protein required for filamentous growth in	
	RT	Ustilago maydis".	
	RL	EMBO J. 16:3464-3473(1997).	
	DL		



Best Local Similarity 39.8%; Pred. No. 1.8e-43;  
Matches 159; Conservative 51; Mismatches 100; Indels 89; Gaps 10;

```
QY      6 VAVCVVRPLNREESLGETAQVYWKTDN--NVLYQVD----- 41
Db      237 VVVCVRNP--SRASSSDSASV--WNCDSKRNRIPTTEHHPAIAKRTTSSERAGAGASIA 293

QY      42 -----GSKSFNFORVFNHNETTKNVYEEIAAPIIDSIAIQYNGTIFAYGQ 86
Db      294 AAPSSHDLHEDPTSTYHFQDKLIITGAQTDDMYHSHIAPVVRAAVEGVNGTVFAYGQ 353

QY      87 TASGKIYTMGSEDHGLGVIPRAIHDIQKIKKPPDRFLLRVSYMETIYNETITDLLCGTQ 146
Db      354 TSGSKTHMSGDAEPGVIPRAVEQIFQMIKDEPDRFLLRVSYLEYNETLKDLLA--- 410

QY      147 KMKPL-----IREDVNRNVYVADLTAEV 171
Db      411 PLPLIATGSGSSIQTTRPASPPIKGSSHAAGSQSCTLRITIEDQKSSRVITGLREEIV 470

QY      172 YTSEMALKMITKGEKSRHYGETKQNRSSRSHTIFRMILESREKGEPSNCEGSKVKVSHLN 231
Db      471 TDANTVLCIQRQDERHVGATDWNERSRSRSHCVFQLTIESRSPAPSASKE--VRISQLN 528

QY      232 LVDLAGSERAAGTGAAGVRLKEGCNINRSLFILGOVIKKLSDGVGG--FINYRDSKLTR 289
Db      529 LIDLASERAA---SQAERRKGAFAFNKSLTLTGTIGKLTPEVNGDAHIPYRDSKLTR 585

QY      290 ILQNSLGNPKTRIICTITPVS--FDETLTALOFATAK 326
Db      586 ILQTSLSGNARIIVICTLSFDPTEHANETLSTLAFGRCK 624
```

Search completed: July 29, 2004, 09:39:39  
Job time : 14.9321 secs

Black sheet  
vs P10